

2-ASIS LEIDIMAS  
ATNAUJINTAS IR  
IŠPLĒSTAS

# DATA DRIVEN CONSTRUCTION

NAVIGACIJA STATYBŲ PRAMONĖJE  
INFORMACIJOS AMŽIUJE

DIRBTINIO INTELEKTO IR  
LLM NAUDOJIMO ATVEJŲ  
NAUDOJIMAS

Artem Boiko



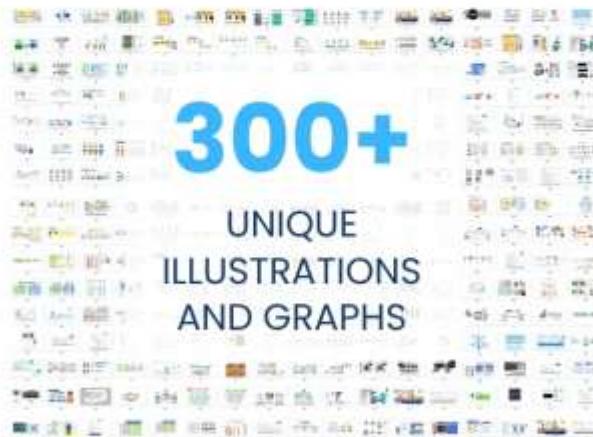
100+

KEY DATA-  
RELATED  
TOPICS



50+

PRACTICAL  
PROBLEMS  
SOLVED



300+

UNIQUE  
ILLUSTRATIONS  
AND GRAPHS

# DATA-DRIVEN CONSTRUCTION

Navigācija datu laikmetā  
būvniecības nozarē

Otrais, pārstrādāts un papildināts izdevums

ARTEM BOIKO

**“**

"Boiko ir IT Džeimss Karvils (James Carville of IT) - viņa daudzkārt citētajā "It's the economy, stupid" ("Tā ir ekonomika, muļķi!") ir tikai viens vārds, kas jānomaina pret šo slaveno grāmatu. "Tie ir dati, muļķi." (Un, lai atrastu savu ceļu datu visumā, vēl šodien ir spēkā seno romiešu teiciens, kas aizgāja no grieķu valodas: "Navigare necesse est". Autors ar drošu roku un nelokāmu kompasu vada savus lasītājus cauri visiem datu okeāna dzīlēm un seklumiem, nemaz nerunājot par vispusīgu vēsturisko pieeju un, visbeidzot, ļoti oriģinālu grafiku un labu humora izjūtu, kas nav pamanāma tikai no otrā acu uzmetiena. Starptautiskā reakcija uz Boiko grāmatu svārstās no eiforiskas atzinības līdz diezgan žēlīgam skepticismam, kas ir noderējis grāmatas otrajam vācu valodas izdevumam. Boiko ir oriģināls un nedogmatisks datu domātājs. Viņš lasītājam piedāvā aizraujošas atziņas un vienmēr drosmīgas, pat provokatīvas tēzes, kas iedvesmo tālākai domāšanai. Lieliskas zāles pret vācu latentā konsensualisma slimību. Starp citu, iepriekš minētajam latīņu sakāmvārdam ir papildinājums: "vivere non est necesse". Tā neattiecas uz Boiko pieeju datu pasaulei - dati dzīvo, un to dzīve ir nepieciešama, lai neteiktu izšķiroša".

- Dr. Burkhard Talebitari, ārštata redaktors, tostarp žurnāla:  
BIM, ko kopš 2013. gada ik gadu izdod Ernst & Sohn.

"Artema Boiko grāmata ir pagrieziena punkts digitalizācijas demokratizācijā būvniecības nozarē - un tā ir īsts spēles pavērsiens mazajiem un vidējiem uzņēmumiem (MVU). Īpaši revolucionārs: izmantojot modernus atvērtā koda (open-source low-code un no-code) rīkus, uzņēmumi jau tagad var efektīvi integrēt datus savos biznesa procesos un rentabli tos analizēt - bez padzījinātām programmēšanas zināšanām. Tādējādi kļūst lieki izmantot dārgas un apgrūtinošas komerciālās programmatūras paketes. Šī grāmata ir aicinājums rīkoties! Tā ir vērtīgs ceļvedis ikvienam, kurš ne tikai vēlas izprast būvniecības nozares digitālo transformāciju, bet arī vēlas to aktīvi veidot - pragmatiski, efektīvi un tālredzīgi. Tagad ir pienācis laiks sadarbīties, lai daļtos ar šīm zināšanām un ilgtspējīgi palielinātu būvniecības nozares produktivitāti."

- **Dr. Maiks Makss Bēlers (Michael Max Buehler)**, HTWG Konstancē būvniecības vadības profesors, GemeinWerk Ventures līdzīpašnieks un DevvStream neatkarīgais direktors.

"DataDrivenConstruction" grāmata ir viens no pirmajiem soljiem ārpus ierastās būvnieku pasaules robežām ar sarežģītajām projektaešanas un vadības sistēmām, kad, šķiet, datu sarežģītība un piesātinātība pat nedod iespēju radikāli vienkāršot un palielināt darba ar būvniecības datiem pārredzamību. Savā grāmatā Artjoms vienkāršā valodā parāda, kādas iespējas mūsdienīgas tehnoloģijas darbam ar datiem paver mūsu priekšā, un burtiski sniedz konkrētus soļus, ko var nekavējoties pielietot savā darbā. Aicinu ikvienu, kurš vēlas saprast, kurp virzīsies automatizācijas sistēmas būvniecības nozarē, rūpīgi izpētīt šo grāmatu, lai saprastu, ka datu revolūcija būvniecībā jau kļauvē pie mūsu durvīm. Tagad tā interesē tikai ģeķus, bet pēc dažiem gadiem, tāpat kā BIM, šādas pieejas un programmatūra būs visur!"

- **Ihor Rogachew**, RGD IMT kompetences centra, BIM un digitālās transformācijas vadītājs un InfraBIM.Pro dibinātājs.

"Es Joti iesaku grāmatu DataDrivenConstruction, kurā, kā teikts nosaukumā, ir aplūkota uz datiem balstīta informācijas pārvaldības pieeja AECO. Pašlaik to izmantoju, lai paīdzētu uzsākt vairākas diskusijas ar dažādām grupām. Es to uzskatu par Joti pieejamu uzziņu. Papildus rūpīgam pārskatam par AECO rīku vēstures kontekstu, datiem un vairāku galveno tehnoloģiju ievisēšanai grāmatā ir iekļautas vairākas Joti noderīgas diagrammas, kas ieskicē datu avotu un galalietotāju artefaktu darbības jomu, kā arī darba plūsmu paraugus. Man šķiet, ka tieši šāda veida diagrammas mums ir vairāk vajadzīgas, izstrādājot un uzraudzīt informācijas stratēģijas un veicinot BEP - definējot vispārējo uzņēmuma datu modeli, uz kura var pārkāpt PIM un AIM robežu."

- **Pols Ranslijs (Paul Ransley)**, Acmena galvenais konsultants un sistēmu integrācijas inženieris Londonas transporta uzņēmumā.

"Ja "dati ir jaunā nafta", mums ir jāiemācās tos definēt, atrast, iegūt, pilnveidot, lai padarītu tos vērtīgus. Man grāmata "DataDrivenConstruction" ir Joti informatīva un izzinoša. Grāmata ir sniepts noderīgs vēsturiskais fons un vienkāršā valodā izskaidrota darbošanās ar datiem. Tiem, kurus interesē digitālā transformācija, tā sniedz labu izpratni par datiem - kā tie darbojas, kā tie ir strukturēti un kā tos var izmantot."

- **Ralph Montague**, ArcDox direktors, BIM koordinatoru samita direktors un Īrijas Nacionālās Standartizācijas iestādes BIM Nacionālās spoguļkomitejas priekšsēdētājs.

"Kā uzsvērts grāmatā, informācija būvniecības nozarē ir būtiska vērtība, un tās pieejamība pieejamos formātos ievērojami atvieglo precīzu lēmumu pieņemšanu un paātrina projektu izpildes termiņus. Grāmata piedāvā neitrālu un efektīvu pieeju, kā piekļūt šim avotam un izmantot to lēmumu pieņemšanā. Grāmatā izklāstīta metodoloģija izmanto mūsdienīgu pieeju, kas apvieno mākslīgā intelekta vadītu programmēšanu ar pieejamiem atvērtā koda rīkiem. Izmantojot mākslīgā intelekta jaudu un izmantojot atvērtā pirkoda programmatūru, metodoloģijas mērķis ir uzlabot automatizāciju, optimizēt procesus un veicināt pieejamību un sadarbību šajā jomā. Grāmatas valoda ir skaidra un viegli uzverama."

- **Dr. Salih Ofluoğlu**, Antālijas Bilim Universitātes Tēlotājmākslas un arhitektūras fakultātes dekāns un Eirāzijas BIM foruma organizators.

"Viss, ko es varu teikt, ir WOW! Veids, kādā jūs iekļāvāt vēsturi, LLM, grafikas un vispārējā vieglums saprast jūsu punktus, ir patiesi ievērojams. Grāmatas plūsma ir pārsteidzoša. Šajā grāmatā ir tik daudz izcilu aspektu; tā patiesi maina spēles noteikumus. Tas ir lielisks informācijas avots, un es izsaku jums atzinību par pūlēm un aizrautību, ko esat tajā ieguldījis. Apsveicu jūs ar tik ievērojama darba radīšanu. Es varētu turpināt, bet pie tiek pateikt, ka esmu neticami pārsteigts!"

- **Natasha Prinsloo**, energylab\_ digitālās prakses vadītāja.

"Ikvienam būvniecības nozarē strādājošajam, sākot no iesācēja līdz pieredzējušam profesionālim, šī grāmata mainīs spēles gaitu! Tā nav tipiska putekļaina lasāmviela - tajā ir daudz ieskatu, stratēģiju un humora, kas neļaus jums aizrauties. Grāmata aptver datu izmantošanas attīstību būvniecībā - no senām datu reģistrēšanas metodēm līdz jaunākajām digitālajām tehnoloģijām. Tas ir kā ceļojums laika mašīnā cauri būvniecības datu evolūcijai. Neatkarīgi no tā, vai esat arhitekts, inženieris, projektu vadītājs vai datu analītikis, šīs visaptverošais ceļvedis revolucionāri mainīs jūsu pieeju projektiem. Esiet gatavi optimizēt procesus, uzlabot lēmumu pieņemšanu un vadīt projektus kā nekad agrāk!"

- **Pjērpaolo Vergati (Pierpaolo Vergati)**, Romas Sapienzas Universitātes pasniedzējs un uzņēmuma Fintecna vecākais būvniecības projektu vadītājs.

**"Es izlasīju grāmatu vienā elpas vilcienā, mazāk nekā 6 stundās. Grāmatas izgatavošanas kvalitāte ir lieliska, biezs glancēts papīrs, krāsu gamma, patīkams fonts. Lielais praktisko piemēru skaits, kā strādāt ar LLM, kas raksturīgi būvniecības nozarei, ietaupīs jums mēnešiem, ja ne gadiem ilgas pašmācības. Darba piemēri ir ļoti daudzveidīgi, sākot no vienkāršiem un beidzot ar sarežģītiem, neprasot iegādāties sarežģītu un dārgu programmatūru. Grāmata ļaus jebkura būvniecības nozares uzņēmuma īpašniekiem no jauna paraudzīties uz savu biznesa stratēģiju, digitalizāciju un attīstības perspektīvām. Savukārt mazākiem uzņēmumiem - palielināt efektivitāti, izmantojot pieejamus un bezmaksas rīkus."**

- **Mihails Kosarevs**, TIM-ASG pasniedzējs un konsultants digitālās transformācijas jautājumos būvniecības nozarē.

**Grāmata "DATA DRIVEN CONSTRUCTION"** maina spēles noteikumus ikvienam, kurš interesējas par to, kurp virzās būvniecības nozare datu laikmetā. Artems ne tikai ieskrāpē virsmu; viņš iedzīlinās pašreizējās norisēs, izaicinājumos un daudzsoļošajās iespējās būvniecībā. Šī grāmata izceļas ar savu pieejamību - Artems sarežģītas idejas skaidro, izmantojot saistošas analogijas, kas padara saturu viegli uztveramu. Man šī grāmata šķita neticami informatīva un vienlaikus saistoša. Kopumā Artems ir radījis vērtīgu resursu, kas ne tikai informē, bet arī iedvesmo. Neatkarīgi no tā, vai esat pieredzējis profesionālis vai iesācējs būvniecības nozarē, šī grāmata paplašinās jūsu redzesloku un padzīlinās izpratni par to, kurp šī nozare virzās. Ľoti iesakām!"

- **Moayad Saleh**, arhitekts un BIM ieviešanas vadītājs uzņēmumā TMM GROUP Gesamtplanungs GmbH.

"Jāteic, ka "Data-Driven Construction" ir grāmata, ko ir vērts mācīt kā mācību grāmatu universitātēs, un tā būs vērtīgs ieguldījums BIM jomā, kas attīstās. Data-Driven Construction grāmatā ir iekļauts tehniskais glosārijs, kas ļoti labi izskaidro jēdzienus. Tēmas, kuras ir ārkārtīgi grūti izskaidrot, ir padarītas vienāršas un saprotamas ar ļoti skaitu vizuālu valodu. Domāju, ka tas, ko paredzēts skaidrot vizuāli, lasītājam ir jāizsaka, kaut vai īsi. Dažu vizuālo materiālu saprotamība, citiem vārdiem sakot, lasot vizuālo materiālu, ir nepieciešama atsevišķa informācija. Vēlos arī teikt, ka ar Artema Boiko vērtīgo darbu labprāt iepazīstinu savās lekcijās un semināros augstskolās."

- **Dr. Ediz Yazıcıoglu**, ArchCube īpašnieks un Būvniecības projektu vadības pasniedzējs Stambulas Tehniskās universitātes Arhitektūras fakultātē un Medipol Universitātē.

**"Uz datiem balstīta būvniecība"** spilgti atspoguļo uz informāciju balstīta darba ar būvniecības datiem pamatus. Grāmata, kurā aplūkotas informācijas plūsmas un ekonomikas pamatjēdzieni, tādējādi izcejas uz citu BIM grāmatu fona, jo tā ne tikai pārstāv programmatūras ražotāja perspektīvu, bet arī cenšas nodot pamatjēdzienus. Grāmata, ko vērts izlasīt un redzēt."

- **Jakob Hirn**, Build Informed GmbH izpilddirektors un līdzdinātājs, inovāciju foruma "On Top With BIM" iniciators.

**"Dati ir jaunā nafta"**, kā mēdz teikt, tāpēc tās meklētājiem vai kalnračiem ir jābūt pareizajiem rīkiem un domāšanas veidam, lai iegūtu vērtību no šī 21. gadsimta resursa. Būvniecības nozare pārāk ilgi ir atradusies uz "3D informācijas" balstītu procesu sliedēs, kad projekta realizācija balstās uz kāda cita uzceptu informāciju (piemēram, viņi jau ir uzzīmējuši pīrāgu vai joslu diagrammu), lai gan pamatā esošie "dati" (piemēram, neapstrādāta izklājlapa) spēj sniegt daudz vairāk, jo īpaši tāpēc, ka daudzu datu saplūšana un mākslīgais intelekts sniedz neierobežotas iespējas. Ja sniedzat (vai mācāt/veicat pētījumus) būvniecības jomā, šī grāmata ir jūsu labākais - un pagaidām vienīgais - resurss, lai orientētos uz datiem balstītajā pasaulē, kurā esam nonākuši."

- **Dr. Zulfikar Adamu**, LSBU asociētais profesors stratēģiskās IT būvniecības jomā, Apvienotā Karaliste.

Artema Boiko grāmata "Data-Driven Construction" ir iespaidīgs darbs, kas piedāvā stabili pamatu būvniecības nozarei laikā, kad nepārtraukti attīstās tehnoloģijas un informācijas iespējas. Boiko izdodas saprotamā veidā izklāstīt sarežģītas tēmas, vienārku ieviešot arī vizionāras idejas. Grāmata ir pārdomāts apkopojums, kurā ne tikai izceertas pašreizējās norises, bet arī sniegs skatījums uz nākotnes inovācijām. Grāmata ir ļoti ieteicama ikvienam, kas vēlas apgūt uz datiem balstītu būvniecības plānošanu un izpildi."

- **Markus Eiberger**, Štutgartes Lietišķo zinātņu universitātes pasniedzējs, Projekta vadītājs un filiāles vadītāja vietnieks uzņēmumā Konstruktionsgruppe Bauen, BIM klastera Bädene-Virtemberga asociācijas valdes loceklis.



Otrais izdevums, 2025. gada marts.  
© 2025 | Artem Boiko | Karlsruhe

ISBN 978-3-912002-09-6



Artem Boiko Autortiesības

[boikoartem@gmail.com](mailto:boikoartem@gmail.com)  
[info@datadrivenconstruction.io](mailto:info@datadrivenconstruction.io)

Nevienu šīs grāmatas daļu nedrīkst reproducēt vai pārraidīt jebkādā formā vai ar jebkādiem elektroniskiem vai mehāniskiem līdzekļiem, ieskaitot fotokopēšanu, ierakstīšanu vai jebkādu informācijas glabāšanas un izgūšanas sistēmu, bez autora rakstiskas atļaujas, izņemot nekomerciālu izplatīšanu nemainītā formā. Grāmata tiek izplatīta bez maksas, un to var brīvi kopīgot ar citiem lietotājiem personīgiem, izglītības vai pētniecības mērķiem, ja tiek saglabāta autorība un atsauces uz oriģinālu. Autors saglabā visas nepatentētās tiesības uz tekstu un nesniedz nekādas tiešas vai netiešas garantijas. Grāmatā minētie uzņēmumi, produkti un nosaukumi var būt izdomāti vai izmantoti kā piemēri. Autors nav atbildīgs par jebkādām sniegtās informācijas izmantošanas sekām. Grāmatā ietvertā informācija tiek sniegtā "tāda, kāda tā ir", bez pilnīguma vai atbilstības garantijas. Autors nav atbildīgs par nejaušiem vai izrietošiem zaudējumiem, kas radušies, izmantojot šajā grāmatā iekļauto informāciju, kodu vai programmas. Šajā grāmatā sniegtie koda paraugi ir paredzēti tikai izglītojošiem nolūkiem. Lasītāji tos izmanto uz savu risku. Autors iesaka visus programmatūras risinājumus pirms lietošanas rāzošanas vidē pārbaudīt. Visas tekstā minētās preču zīmes un produktu nosaukumi ir attiecīgo uzņēmumu preču zīmes, reģistrētas preču zīmes vai pakalpojumu zīmes, un tie ir attiecīgo īpašnieku īpašums. Šo nosaukumu izmantošana grāmatā nenozīmē, ka tie ir saistīti ar to īpašniekiem vai ka tie tos atbalsta. Trešo pušu produktu vai pakalpojumu pieminēšana nav uzskatāma par ieteikumu vai apstiprinājumu. Piemēros izmantotie uzņēmumu un produktu nosaukumi var būt to īpašnieku preču zīmes. Saites uz trešo personu tīmekļa vietnēm ir sniegtas ērtības labad, un tās nenozīmē, ka autors atbalsta šajās vietnēs sniegtā informāciju. Visi minētie statistikas dati, citāti un pētījumi bija aktuāli rakstīšanas laikā. Laika gaitā dati var mainīties.

Šī grāmata tiek izplatīta saskaņā ar Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licenci (CC BY-NC-ND 4.0). Jūs varat to kopēt un izplatīt nekomerciālos nolūkos ar nosacījumu, ka saglabājat autortiesības un neizdara nekādas izmaiņas.



CC BY-NC-ND 4.0

© 2024 Artjoms Boiko. Pirmais izdevums.  
© 2025 Artjoms Boiko. Otrais, pārstrādāts un papildināts izdevums.  
Visas tiesības aizsargātas.

## PRIEKŠVĀRDS OTRAJAM IZDEVUMAM

Šī grāmata ir rezultāts dzīvam dialogam ar profesionālu kopienu. Tā ir balstīta uz daudzām profesionālām diskusijām par datu pārvaldību būvniecības nozarē, kas notika dažādās profesionālās platformās un sociālajos medijs. Šīs diskusijas kļuva par pamatu rakstiem, publikācijām un vizuālajiem materiāliem, kas izraisīja plašu rezonansu profesionālajā sabiedrībā. Autora saturs katru gadu piesaista miljoniem skatījumu dažādās platformās un valodās, apvienojot būvniecības digitalizācijas jomas profesionālus.

Gada laikā pēc pirmā izdevuma publicēšanas grāmatu pasūtīja eksperti no vairāk nekā 50 valstīm - no Brazīlijas un Peru līdz Maurīcijai un Japānai. Grāmatas otrs izdevums, kuru jūs tagad turat rokās, ir pārstrādāts un papildināts, pamatojoties uz ekspertu atsauksmēm, pirmā izdevuma kritiku un diskusijām profesionālu aprindās. Pateicoties atsauksmēm, otrs izdevums ir būtiski papildināts: pievienotas jaunas nodajas par CAD (BIM) tehnoloģijām un efektīvu ETL -procesu izveidi. Ievērojami palielināts arī praktisko piemēru un gadījumu analīzes skaits. Īpaši vērtīgas ir atsauksmes no būvniecības nozares vadītājiem, konsultāciju uzņēmumiem un lielākajiem IT uzņēmumiem, kuri vērsās pie autora ar jautājumiem par digitalizāciju un sadarbspēju gan pirms, gan pēc grāmatas pirmās versijas publicēšanas. Daudzi no tiem jau ir izmantojuši grāmatā aprakstītās pieejas vai plāno to darīt tuvākajā nākotnē

Jūsu rokās ir grāmata, kas tapusi diskusiju un aktīvas viedokļu apmaiņas rezultātā. Progress dzimst dialogā, viedokļu sadursmē un atvērtībā jaunām pieejām. Paldies, ka piedalāties šajā dialogā. Jūsu konstruktīvā kritika ir pamats turpmākajiem uzlabojumiem. Ja tekstā atrodat kļūdas vai vēlaties dalīties ar idejām un ieteikumiem, gaidīsim jebkuru atsauksmi. Kontaktinformācija ir norādīta grāmatas beigās

## KĀPĒC GRĀMATA IR BEZMAKSAS?

Šī grāmata tika iecerēta kā atvērts izglītības resurss, kura mērķis ir izplatīt mūsdienīgas pieejas datu pārvaldībai būvniecības nozarē. Grāmatas pirmā versija kalpoja par pamatu profesionālu kopienas komentāru un ieteikumu apkopošanai, kas ļāva uzlabot materiāla struktūru un saturu. Visi komentāri, ieteikumi un idejas ir rūpīgi analizēti un iekļauti šajā pārstrādātajā versijā. Grāmatas mērķis ir palīdzēt būvniecības profesionāļiem saprast, cik svarīgi ir strādāt ar datiem: sistemātiski, apzināti un domājot par informācijas ilgtermiņa vērtību. Autors ir apkopojis piemērus, ilustrācijas un praktiskus novērojumus, kas gūti, vairāk nekā 10 gadus strādājot būvniecības digitalizācijas jomā. Lielākā daļa šī materiāla radusies, ištekojot reālus projektus, diskutējot ar inženieriem un būvniekiem, piedaloties starptautiskās iniciatīvās un mācību semināros. Grāmata ir mēģinājums strukturēt uzkrāto pieredzi un dalīties ar to pieejamā veidā. Ja vēlaties atbalstīt grāmatas ideju tālāku izplatīšanu un iegūt ērtu formātu lasīšanai, darbam ar piemēriem un vizuālajiem materiāliem - varat iegādāties [drukāto versiju](#).

## IZMANTOŠANAS PRATIESĪBAS

Visus šīs grāmatas materiālus, ilustrācijas un fragmentus var reproducēt, citēt vai izmantot jebkurā formātā un jebkurā informācijas nesējā ar nosacījumu, ka tiek norādīts avots: Artem Boiko autorība un grāmatas nosaukums "Data-Driven Construction". Paldies par cieņu pret darbu un zināšanu izplatīšanu.

Ar patiesu pateicību šo grāmatu savai ģimenei, kas man jau no  
mazotnes ieaudzināja dzīļu mīlestību pret būvniecību, savai  
dzimtai kalnraktuvju pilsētai par izturību un savai sievai  
mērniecei, kuras neatslābstošais atbalsts ir bijis mans  
pastāvīgais iedvesmas avots.

## KAM ŠĪ GRĀMATA IR PAREDZĒTA

Šī grāmata, kas uzrakstīta saprotamā valodā, ir paredzēta plašam būvniecības nozares lasītāju lokam - sākot no studentiem un iesācējiem, kuri vēlas apgūt mūsdienu būvniecības procesu pamatus, līdz profesionāliem, kuriem nepieciešama mūsdienīga metodoloģija datu pārvaldībai būvniecībā. Neatkarīgi no tā, vai esat arhitekts, inženieris, būvuzraugs, būvdarbu vadītājs vai datu analītikis, šīs visaptverošais ceļvedis ar daudzām unikālām ilustrācijām un grafikiem piedāvā vērtīgu ieskatu par to, kā izmantot datus uzņēmējdarbībā, lai optimizētu un automatizētu procesus, uzlabotu lēmumu pieņemšanu un vadītu būvniecības projektus dažādos līmenos, izmantojot modernus rīkus.

Grāmata ir visaptverošs ceļvedis, kurā apvienoti teorētiskie pamati un praktiski ieteikumi datu pārvaldības metožu integrēšanai būvniecības procesos. Grāmatā galvenā uzmanība pievērsta stratēģiskai informācijas izmantošanai, lai optimizētu darbību, automatizētu procesus, uzlabotu lēmumu pieņemšanu un efektīvi pārvaldītu projektus, izmantojot modernus digitālos rīkus.

Šajā grāmatā aplūkoti teorētiskie un praktiskie aspekti darbam ar informāciju būvniecības nozarē. Ar detalizētu piemēru palīdzību tajā aplūkota uzdevumu parametru noteikšanas metodoloģija, prasību apkopošana, nestukturētu un daudzformātu datu apstrāde un to pārveidošana efektīvos risinājumos būvniecības uzņēmumiem.

Lasītājs secīgi iziet ceļu no prasību veidošanas un datu pamatmodelju izstrādes līdz sarežģītākiem procesiem - heterogēnu informācijas avotu integrācijai, ETL -procesu izveidei, informācijas cauruļvadu un mašīnmācīšanās modeļu izveidei. Secīgā pieeja ļauj uzskatāmi parādīt biznesa procesu un lēmumu atbalsta sistēmu organizācijas un automatizācijas mehānismus būvniecības nozarē. Katrā grāmatas daļa noslēdzas ar praktisku nodalju, kurā ietverti soli pa solim izstrādāti norādījumi, kas ļauj iegūtās zināšanas nekavējoties pielietot reālos projektos.

# GRĀMATAS DAĻU KOPSAVILKUMS

Šīs grāmatas pamatā ir datu transformācijas koncepcija vērtību kēdē: no datu vākšanas un kvalitātes nodrošināšanas līdz analītiskai apstrādei un vērtīgu praktisku risinājumu iegūšanai, izmantojot mūsdienīgus rīkus un metodoloģijas.

- 1. daļā "Digitālā evolūcija būvniecībā"** ir aprakstīta datu pārvaldības vēsturiskā transformācija no māla plāksnītēm līdz mūsdienu digitālajām sistēmām, analizēta modulāro sistēmu rašanās un pieaugošā informācijas digitalizācijas nozīme rūpniecisko revolūciju kontekstā.
- 2. daļā "Informācijas izaicinājumi būvniecības nozarē"** aplūkotas datu sadrumstalotības, "informācijas silos", HiPPO pieejas ietekmes lēmumu pieņemšanu un patentētu formātu ierobežojumu problēmas, kā arī ierosināts apsvērt pāreju uz mākslīgā intelekta un LLM ekosistēmām.
- 3. daļa: Datu sistematizācija būvniecībā** - veido būvniecības datu tipoloģiju, apraksta to organizācijas metodes, integrāciju ar korporatīvajām sistēmām un aplūko kompetences centru izveidi informācijas procesu standartizācijai.
- 4. daļa: Datu kvalitātes nodrošināšana** - atklāj metodoloģijas, kā pārvērst atšķirīgu informāciju kvalitatīvos, strukturētos datos, tostarp datu ieguvi no dažādiem avotiem, validēšanu un modelēšanu, izmantojot LLM.
- 5. daļā "Izmaksu un laika aprēķini"** aplūkota izmaksu un plānošanas aprēķinu digitalizācija, apjomu iegūšanas automatizācija no CAD (BIM) modeļiem, 4D-8D modelēšanas tehnoloģijas un būvniecības projektu ESG aprēķini.
- 6. daļa: CAD un BIM** - kritiski analizē projektēšanas tehnoloģiju attīstību, sistēmu sadarbības jautājumus, tendences atvērto datu formātu jomā un mākslīgā intelekta izmantošanas perspektīvas projektēšanā.
- 7. daļā "Datu analīze un automatizācija"** aplūkoti informācijas vizualizācijas principi, galvenie veikspējas rādītāji, ETL procesi, darbplūsmas orķestrēšanas rīki un valodas modeļu izmantošana rutīnas uzdevumu automatizācijai.
- 8. daļa: Datu glabāšana un pārvaldība** - tajā aplūkoti datu glabāšanas formāti, datu noliktavas un datu ezera koncepcijas, datu pārvaldības principi un jaunas pieejas, tostarp vektoru datubāzes un DataOps un VectorOps metodoloģijas.
- 9. daļā "Lielie dati un mašīnmācīšanās"** uzmanība pievērsta pārejai uz objektīvu analīzi, kas balstīta uz vēsturiskiem datiem, lietiskais internets būvlaukumos un mašīnmācīšanās algoritmu izmantošana, lai prognozētu projekta izmaksas un termiņus.
- 10. daļā "Būvniecības nozare digitālo datu laikmetā"** aplūkota būvniecības nozares nākotne, analizēta pāreja no cēloņsakarību analīzes uz darbu ar korelācijām, būvniecības "überizācijas" koncepcija un digitālās transformācijas stratēģijas.

# What is meant by **data-driven construction** ?



## IEVADS

Cik ilgi jūsu uzņēmums var saglabāt konkurētspēju pasaulē, kurā tehnoloģijas strauji attīstās un katru uzņēmējdarbības aspektu, sākot ar laika plānošanu un izmaksu aprēķināšanu un beidzot ar riska analīzi, automatizē ar mašīnmācīšanās modeļiem?

Būvniecības nozare, kas pastāv tikpat ilgi kā pati cilvēce, atrodas uz revolucionāru pārmaiņu sliekšņa, kas sola pilnībā mainīt mūsu priekšstatus par tradicionālo būvniecību. Jau citās tautsaimniecības nozarēs digitalizācija ne tikai maina noteikumus, bet arī nežēlīgi izspiež no tirgus uzņēmumus, kuri nav spējuši pielāgoties jaunajai datu apstrādes videi un nespēj uzlabot lēmumu pieņemšanas ātrumu (1. attēls).



1. attēls Lēmumu pieņemšanas ātrums būvniecības nozarē biežāk nekā citās nozarēs ir atkarīgs no cilvēciskā faktora.

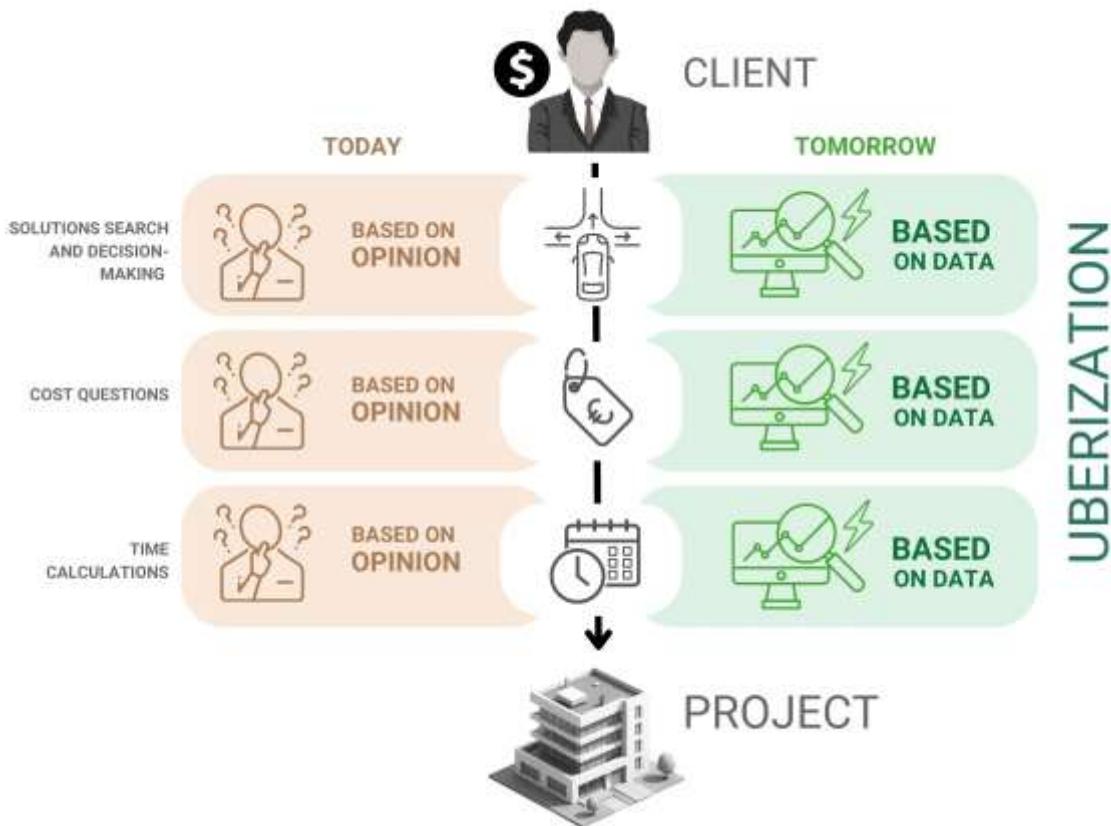
Banku, mazumtirdzniecības, logistikas un lauksaimniecības nozares strauji virzās uz pilnīgu digitalizāciju, kur neprecizitātēm un subjektīviem viedokļiem vairs nav vietas. Mūsdienu algoritmi spēj analizēt milzīgus datu apjomus un sniegt klientiem precīzas prognozes - vai tā būtu kredīta atmaksas varbūtība, optimāli piegādes maršuti vai risku prognozēšana.

Būvniecība ir viena no pēdējām nozarēm, kurā nenovēršami notiek pāreja no risinājumiem, kas balstīti uz augsti atalgotu speciālistu viedokļiem, uz risinājumiem, kas balstīti uz datiem. Šo pāreju veicina ne tikai jaunas tehnoloģiskās iespējas, bet arī pieaugošās tirgus un klientu prasības pēc pārredzamības, precīzitātes un ātruma.

Robotizācija, procesu automatizācija, atvērtie dati un uz tiem balstītas prognozes - tas viss vairs nav tikai iespēja, bet gan neizbēgamība. Lielākā daļa būvniecības nozares uzņēmumu, kas vēl nesen bija atbildīgi klientam par projektu apjoma, izmaksu, laika un kvalitātes kontroles aprēķināšanu, tagad riskē

kļūt tikai par pasūtījumu izpildītājiem, nevis pieņemt galvenos lēmumus (2. attēls).

Attīstoties skaitļošanas jaudai, mašīnmācīšanās algoritmiem un datu piekļuves demokratizācijai, tagad ir iespējams automātiski apvienot datus no dažādiem avotiem, kas ļauj padziļināti analizēt procesus, prognozēt riskus un optimizēt izmaksas būvniecības projekta apspriešanas posmos. Šīm tehnoloģijām ir potenciāls radikāli uzlabot efektivitāti un samazināt izmaksas visā nozarē.



2. attēls Klients nav ieinteresēts pārmērīgā cilvēkfaktoru izmantošanā sava projekta realizācijas gaitā.

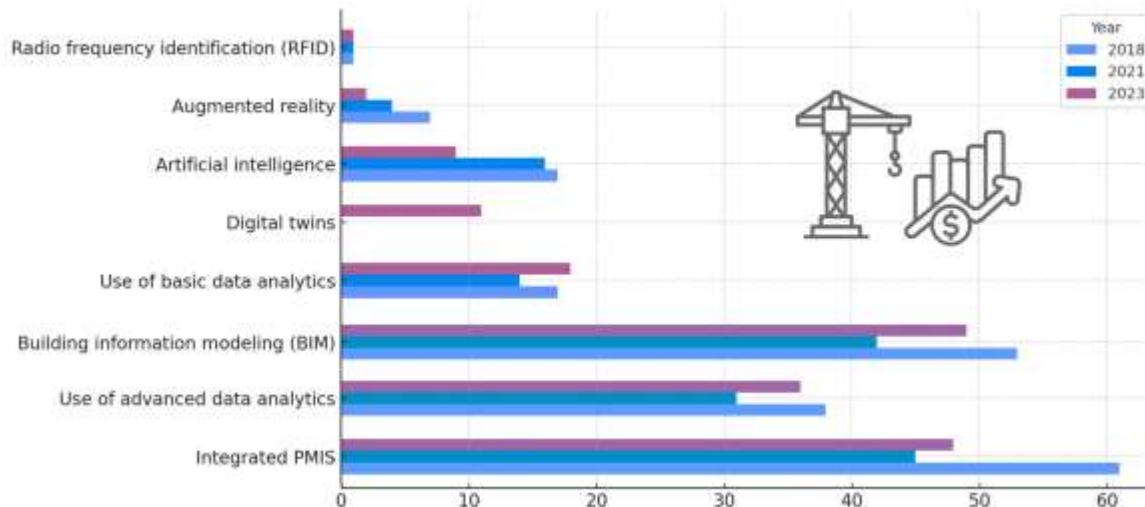
Neraugoties uz visu jauno rīku un koncepciju priekšrocībām, būvniecības nozare jauno tehnoloģiju ieviešanā ievērojami atpaliek no citām tautsaimniecības nozarēm.

Saskaņā ar IT Metrics Key Data 2017 ziņojumu būvniecības nozare IT izdevumu ziņā ieņem pēdējo vietu starp 19 citām nozarēm [1].

Straujais datu apjoma un procesu sarežģītības pieaugums kļūst par uzņēmumu vadības galvassāpēm, un galvenā problēma jauno tehnoloģiju izmantošanā ir tā, ka dati, neraugoties uz to pārpilnību, joprojām ir sadrumstaloti, nestrukturēti un bieži nesaderīgi starp dažādām sistēmām un programmatūras produktiem. Tāpēc daudzus būvniecības nozares uzņēmumus šobrīd galvenokārt uztrauc datu kvalitātes jautājumi, kurus var atrisināt tikai ieviešot efektīvas, automatizētas pārvaldības un analītikas sistēmas.

Saskaņā ar KPMG® 2023. gadā veikto būvniecības vadītāju aptauju [2], projektu vadības informācijas sistēmām (PMIS), uzlabotai un pamata datu analīzei un būves informācijas modelēšanai (BIM) ir vislielākais potenciāls uzlabot projektu INI (3. attēls).

Technologies with Potential to Deliver the Greatest Overall ROI



**3. attēls Būvniecības uzņēmumu vadītāju aptauja: kuras tehnoloģijas nodrošinās vislielāko ieguldījumu atdevi (ROI) kapitālprojektos? (pamatojoties uz materiāliem [2]).**

Risinājums izaicinājumiem, kas saistīti ar datu integrēšanu uzņēmējdarbības procesos, ir nodrošināt augstas kvalitātes informāciju, izmantot piemērotus datu formātus un piemērot efektīvas metodes datu radīšanai, glabāšanai, analīzei un apstrādei.

Datu vērtības apzināšanās liek dažādām nozarēm atteikties no izolētām lietojumprogrammām un sarežģītām birokrātiskām pārvaldības struktūrām. Tā vietā uzmanība tiek pievērsta jaunu pieeju radīšanai informācijas arhitektūrai, pārveidojot uzņēmumus par mūsdienīgiem, uz datiem balstītiem uzņēmumiem. Agri vai vēlu arī būvniecības nozare pati spers šo soli, pārejot no pakāpeniskas digitālās evolūcijas uz īstu digitālo revolūciju, kas skars visus uzņēmumus.

Pāreja uz uz datiem balstītiem uzņēmējdarbības procesiem nebūs viegla. Daudzi uzņēmumi saskarsies ar grūtībām, jo vadītāji ne vienmēr saprot, kā izmantot haotiskas datu kopas, lai uzlabotu efektivitāti un biznesa izaugsmi.

Šī grāmata iedzījinās datu pasaulē, kur informācija kļūst par galveno stratēģisko resursu, kas nosaka uzņēmējdarbības procesu efektivitāti un ilgtspēju. Strauji pieaugot informācijas apjomam, uzņēmumi saskaras ar jauniem izaicinājumiem. Digitālā transformācija vairs nav tikai modes vārds - tā kļūst par nepieciešamību.



4. attēls Datī un procesi ir būvniecības pamats.

Izprast transformāciju nozīmē spēt izskaidrot sarežģīto vienkāršiem vārdiem. Tāpēc grāmata ir uzrakstīta saprotamā valodā un papildināta ar autora ilustrācijām, kas īpaši radītas, lai skaidri izskaidrotu galvenos jēdzienus. Šīs diagrammas, shēmas un vizualizācijas ir veidotas tā, lai likvidētu uztveres šķēršļus un padarītu materiālu saprotamu pat tiem, kas iepriekš uzskatīja šādas tēmas par pārāk sarežģītām. Visas šajā grāmatā iekļautās ilustrācijas, diagrammas un grafikas ir autora veidotas un īpaši izstrādātas, lai vizualizētu galvenos tekstā aprakstītos jēdzienus.

*Viens attēls ir tūkstoš vārdu vērts [3].*

- Fred R. Barnard, angļu ilustrators, 1927. gads.

Lai teoriju sasaistītu ar praksi, mēs izmantosim mākslīgā intelekta rīkus (jo īpaši valodas modeļus), kas Jauj izstrādāt risinājumus bez dzīlām programmēšanas zināšanām. Ja esat orientēts uz praktisku materiālu un jūs vairāk interesē praktisks darbs ar datiem, varat izlaist pirmo ievaddaļu un doties uzreiz uz grāmatas otro daļu, kur sākas konkrētu piemēru un gadījumu apraksts.

Tomēr nelieciet pārāk lielas cerības uz mākslīgā intelekta (AI), mašīnmācīšanās un LLM (Large Language Models) rīkiem kopumā. Bez kvalitatīviem ievades datiem un padziļinātas izpratnes par atiecīgo tematu pat vismodernākie algoritmi nevar sniegt ticamus un jēgpilnus rezultātus.

Microsoft izpilddirektors Satja Nadella (Satya Nadella) brīdina par mākslīgā intelekta burbuļa risku 2025. gada sākumā [4], salīdzinot pašreizējo ažiotāžu ar dot-com burbuli. Viņš uzsver, ka apgalvojumi par AGI (mākslīgā vispārējā intelekta) atskaites punktu sasniegšanu bez pienācīga pamatojuma ir "bezjēdzīga manipulācija ar rādītājiem". Nadella uzskata, ka mākslīgā intelekta patiesie panākumi būtu jāvērtē pēc tā ieguldījuma pasaules IKP izaugsmē, nevis pēc pārmērīgas uzsvēršanas modes vārdos.

Aiz visiem jaunajām tehnoloģijām un koncepcijām slēpjās sarežģīts un rūpīgs darbs, lai nodrošinātu

datu kvalitāti, uzņēmējdarbības procesu parametrizāciju un rīku pielāgošanu reāliem uzdevumiem.

Uz datiem balstīta pieeja nav produkts, ko var vienkārši lejupielādēt vai iegādāties. Tā ir stratēģija, kas ir jāizstrādā. Tā sākas ar jaunu skatījumu uz esošajiem procesiem un problēmām, un tad ir nepieciešama disciplinēta virzība izvēlētajā virzienā.

Vadošie programmatūras izstrādātāji un lietojumprogrammu pārdevēji nebūs pārmaiņu dzinējspēks būvniecības nozarē, jo daudziem no tiem uz datiem balstīta pieeja apdraud viņu iedibināto uzņēmējdarbības modeli.

*Citas nozares [atšķirībā no būvniecības], piemēram, autobūves nozare, jau ir piedzīvojušas radikālas un grausošas pārmaiņas, un to digitālā transformācija jau ir labi iesākusies. Būvniecības uzņēmumiem ir jārīkojas ātri un izlēmīgi: veikli uzņēmumi gūs milzīgus ieguvumus, savukārt tiem, kas vilcināsies, riski būs nopietni. Padomājiet par apvērsumu, ko šajā nozarē ir izraisījusi digitālā fotogrāfija [5].*

- Pasaules ekonomikas foruma ziņojums "Veidojot būvniecības nākotni", 2016. gads.

Tie uzņēmumi, kas savlaicīgi apzinās jaunās pieejas iespējas un priekšrocības, iegūs ilgtspējīgas konkurences priekšrocības un spēs attīstīties un augt, nebūdami atkarīgi no lielo piegādātāju risinājumiem.

Šī ir jūsu iespēja ne tikai pārvarēt gaidāmo informācijas digitalizācijas vētru, bet arī pārņemt kontroli pār to. Šajā grāmatā jūs atradīsiet ne tikai nozares pašreizējā stāvokļa analīzi, bet arī konkrētus ieteikumus, kā pārdomāt un pārstrukturēt savus procesus un uzņēmumu, lai kļūtu par līderi jaunajā būvniecības laikmetā un uzlabotu savu profesionālo pieredzi.

Būvniecības digitālā nākotne nav saistīta tikai ar jaunu tehnoloģiju un programmu izmantošanu, bet gan ar fundamentālu datu apstrādes un uzņēmējdarbības modeļu pārskatīšanu.

Vai jūsu uzņēmums ir gatavs šīm stratēģiskajām pārmaiņām?



# SATURA RĀDĪTĀJS

<b>IEVADS.....</b>	<b>1</b>
<b>SATURA RĀDĪTĀJS .....</b>	<b>I</b>
<b>II DAĻA NO MĀLA PLĀKSNĒM LĪDZ DIGITĀLAJAI REVOLŪCIJAI: KĀ INFORMĀCIJA IR ATTĪSTĪJUSIES BŪVNIECĪBĀ.....</b>	<b>2</b>
NODAĻA 2.1. DATU IZMANTOŠANAS ATTĪSTĪBA BŪVNIECĪBAS NOZARĒ.....	3
Datu ēras sākums būvniecībā .....	3
No māla un papirusa līdz digitālajām tehnoloģijām.....	4
Process kā rīks uz datiem balstītai pieredzei .....	5
Būvniecības procesa informācijas digitalizācija .....	7
NODAĻA 2.2. TEHNOLOGIJAS UN VADĪBAS SISTĒMAS MŪSDIENU BŪVNIECĪBĀ .....	11
Digitālā revolūcija un moduļu MRP/ERP -sistēmu rašanās .....	11
Datu pārvaldības sistēmas: no datu ieguves līdz uzņēmējdarbības izaicinājumiem .....	13
Uzņēmuma micēlijs: kā dati savienojas ar uzņēmējdarbības procesiem.....	16
NODAĻA 2.3. DIGITĀLĀ REVOLŪCIJA UN DATU EKSPLOZIJA .....	20
Datu buma sākums kā evolūcijas vilnis .....	20
Mūsdieni uzņēmumā ģenerēto datu apjoms.....	22
Datu glabāšanas izmaksas: ekonomiskais aspekts.....	23
Datu uzkrāšanas robežas: no masas līdz nozīmei .....	25
Nākamie solji: no datu teorijas līdz praktiskām pārmaiņām.....	26
<b>III DAĻA KĀ BŪVNIECĪBAS NOZARE SLĪGST DATU HAOSĀ.....</b>	<b>28</b>
NODAĻA 3.1. DATU SADRUMSTALOTĪBA UN "SILOS". .....	29
Jo vairāk rīku, jo efektīvāks bizness?.....	29
Datu atslēgas un to ietekme uz uzņēmuma darbību .....	31
Datu dublēšanās un datu kvalitātes trūkums, kas rodas nesaskaņotības dēļ.....	34
HiPPO jeb viedokļu bīstamība lēmumu pieņemšanā .....	36
Nepārtraukti pieaug uzņēmējdarbības procesu sarežģītība un dinamika. ....	39
Ceturta industriālā revolūcija (Industry 4.0) un piektā industriālā revolūcija (Industry 5.0) būvniecībā.....	42
NODAĻA 3.2. HAOSA PĀRVĒRŠANA KĀRTĪBĀ UN SAREŽĢĪTĪBAS SAMAZINĀŠANA.....	45
Liekais kods un slēgtas sistēmas kā šķērslis produktivitātes uzlabošanai .....	45
No atsevišķām datu noliktavām uz vienotu datu noliktavu .....	47

Integrētas glabāšanas sistēmas Jauj pāriet uz mākslīgā intelekta aģentiem .....	48
No datu vākšanas līdz lēmumu pieņemšanai: celš uz automatizāciju .....	51
Nākamie soļi: pārvērst haosu pārvaldāmā sistēmā.....	53
<b>IV DAĻA DATU SISTĒMA BŪVΝIECĪBAS UZŅĒMĒJDARBĪBAS PROCESOS .....</b>	<b>55</b>
NODAĻA 4.1. DATU TIPI BŪVΝIECĪBĀ.....	56
Svarīgākie datu veidi būvniecības nozarē .....	56
Strukturēti dati .....	60
Relāciju datu bāzes RDBMS un SQL vaicājumu valoda .....	61
SQL - vaicājumi datubāzēs un jaunas tendences .....	64
Nestrukturizēti dati .....	66
Teksta dati: starp nestrukturētu haosu un strukturētu haosu y .....	67
Dalēji strukturēti un brīvi strukturēti dati .....	68
Ģeometriskie dati un to izmantošana .....	69
CAD dati: no projektēšanas līdz datu glabāšanai .....	72
BIM (BOM) koncepcijas rašanās un CAD izmantošana procesos.....	75
NODAĻA 4.2. DATU SASKANOŠANA UN STRUKTURĒŠANA .....	81
Sistēmu aizpildīšana ar datiem būvniecības nozarē.....	81
Datu transformācija: mūsdienu biznesa analīzes kritiskais pamats.....	83
Datu modeli: attiecības datos un attiecības starp elementiem.....	87
Patentēti formāti un to ietekme uz digitālajiem procesiem .....	91
Atvērtie formāti maina pieeju digitalizācijai .....	95
Paradigmas maiņa: atvērtā koda kā programmatūras piegādātāju dominances ēras beigas .....	96
Strukturēti atvērti dati: digitālās transformācijas pamats .....	99
NODAĻA 4.3. LLM UN TO LOMA DATU APSTRĀDĒ UN UZŅĒMĒJDARBĪBAS PROCESOS.....	102
LLM tērzēšanas istabas: ChatGPT, LlaMa, Mistral, Claude, DeepSeek, QWEN, Grok datu procesu automatizēšanai.....	102
Lieli LLM valodas modeļi: kā tas darbojas .....	103
Vietējo LLM izmantošana sensitīviem uzņēmuma datiem .....	106
Pilnīga kontrole pār AI uzņēmumā un kā izvietot savu LLM .....	108
RAG: Intelligent LLM - asistenti ar piekļuvi uzņēmuma datiem. ....	110
NODAĻA 4.4. IDE AR LLM ATBALSTU UN TURPMĀKĀS PROGRAMMĒŠANAS IZMAIŅAS.....	112
IDE izvēle: no LLM eksperimentiem līdz biznesa risinājumiem.....	112
IDE ar LLM atbalstu un turpmākās programmēšanas izmaiņas .....	114
Python Pandas: neaizstājams rīks darbam ar datiem .....	115

DataFrame: universāls tabulāro datu formāts.....	119
Nākamie soļi: ilgtspējīgas datu sistēmas izveide.....	122
<b>V DAĻA DATU KVALITĀTE: ORGANIZĀCIJA, STRUKTURĒŠANA, MODELĒŠANA.....</b>	<b>124</b>
NODAĻA 5.1. DATU KONVERTĒŠANA STRUKTURĒTĀ FORMĀ.....	125
Mācīšanās, kā pārvērst dokumentus, PDF, attēlus un tekstus strukturētos formātos.....	125
Piemērs, kā pārvērst PDF -dokumentu tabulā .....	126
Konvertēt JPEG, PNG attēlu uz strukturētu formu .....	130
Teksta datu pārveidošana strukturētā formā .....	132
CAD datu (BIM) konvertēšana strukturētā formā.....	135
CAD risinājumu piedāvātāji virzās uz strukturētu datu apstrādi.....	140
NODAĻA 5.2. KLASIFIKĀCIJA UN INTEGRĀCIJA: VIENOTA BŪVNIECĪBAS DATU VALODA .....	143
Lēmumu pieņemšanas ātrums ir atkarīgs no datu kvalitātes .....	143
Datu standartizācija un integrācija .....	144
Digitālā sadarbīspēja sākas ar prasībām .....	147
Kopīga būvniecības valoda: klasifikatoru loma digitālajā pārveidē .....	149
Masterformat, OmniClass, Uniclass un CoClass: klasifikācijas sistēmu attīstība .....	152
NODAĻA 5.3. DATU MODELĒŠANA UN IZCILĪBAS CENTRS .....	157
Datu modelēšana: konceptuālais, logiskais un fiziskais modelis .....	157
Praktiska datu modelēšana būvniecības kontekstā .....	161
Datu bāzes izveide, izmantojot LLM .....	163
Datu modelēšanas izcilības centrs (CoE).....	165
NODAĻA 5.4. PRASĪBU SISTEMATIZĀCIJA UN INFORMĀCIJAS APSTIPRINĀŠANA .....	168
Prasību apkopošana un analīze: saziņas pārveidošana strukturētos datos.....	168
Procesu diagrammas un konceptuālo ietvaru efektivitāte .....	172
Strukturētas prasības un RegEx regulārās izteiksmes.....	174
Datu vākšana pārbaudes procesam .....	179
datu verifikācija un verifikācijas rezultāti .....	181
verifikācijas rezultātu vizualizācija .....	186
Datu kvalitātes pārbaužu salīdzinājums ar cilvēka dzīves vajadzībām.....	188
Nākamie soļi: datu pārvēršana precīzos aprēķinos un plānos .....	190
<b>VI DAĻA IZMAKSU UN LAIKA APRĒĶINI: DATU IEKLĀUŠANA BŪVNIECĪBAS PROCESOS.....</b>	<b>192</b>
NODAĻA 6.1. BŪVNIECĪBAS PROJEKTU IZMAKSU APRĒĶINI UN TĀMES .....	193
Būvniecības pamati: daudzuma, izmaksu un laika aplēses.....	193
Projektu paredzamo izmaksu aprēķināšanas metodes.....	194

Uz resursiem balstīta metode izmaksu aprēķināšana un novērtēšana būvniecībā .....	195
Būvniecības resursu datubāze: būvmateriālu un darbu katalogs .....	195
Aprēķinu sastādīšana un darbu izmaksu aprēķināšana, pamatojoties uz resursu bāzi.....	197
Projekta galīgo izmaksu aprēķins: no tāmes līdz budžetam .....	202
<b>NODAĻA 6.2. DAUDZUMA APRĒĶINS UN AUTOMĀTiska TĀMju UN GRAFIku IZVEIDE.....</b>	<b>206</b>
Pāreja no 3D uz 4D un 5D: volumetrisko un kvantitatīvo parametru izmantošana .....	206
5D atribūti un atribūtu apjomu iegūšana no CAD.....	206
QTO Quantity Take-Off: projekta datu grupēšana pēc atribūtiem .....	210
QTO automatizācija, izmantojot LLM un strukturētus datus .....	215
QTO visa projekta aprēķins, izmantojot grupas noteikumus no Excel izklājlapas. ....	219
<b>NODAĻA 6.3. 4D, 6D -8D UN OGLEKĻa DIOKSīDA EMISIJu APRĒĶINS CO<sub>2</sub>. ....</b>	<b>225</b>
4D modelis: laika integrēšana būvniecības aplēsēs .....	225
Būvdarbu grafiks un tā automatizācija, pamatojoties uz izmaksu aprēķina datiem .....	226
Paplašinātie atribūtu slāni 6D -8D: no energoefektivitātes līdz drošības nodrošināšanai.....	228
CO <sub>2</sub> aplēses un oglekļa dioksīda emisiju aprēķins būvniecības projektos .....	231
<b>NODAĻA 6.4. BŪVniecības ERP UN PMIS SISTĒMAS .....</b>	<b>235</b>
Būvniecības ERP -sistēmas uz aprēķinu un tāmju piemēra.....	235
PMIS: starpnieks starp ERP un būvlaukumu. ....	240
Spekulācijas, peļņa, izolētība un trūkums pārredzamība ERP un PMIS .....	241
Slēgtas ERP /PMIS ēras beigas: būvniecības nozarei nepieciešamas jaunas pieejas. ....	243
Nākamie soļi: efektīva projektu datu izmantošana.....	245
<b>VII DAĻA CAD UN BIM: MĀRKETINGS, REALITĀTE UN PROJEKTĒŠANAS DATU NĀKOTNE BŪVniecībā 248</b>	
<b>NODAĻA 7.1. BIM- KONCEPCIJU PARĀDīŠANĀS BŪVniecības NOZARĒ .....</b>	<b>249</b>
BIM un atvērtā BIM kā CAD- piegādātāju mārketinga koncepciju rašanās vēsture. ....	249
BIM realitāte: integrētu datu bāzu vietā slēgtas moduļu sistēmas .....	252
Atvērtā formāta IFC parādīšanās būvniecības nozarē.....	254
IFC formāta problēma atkarībā no ģeometriskā kodola .....	256
Parādīšanās semantikas tēmas konstrukcijā un ontoloģijā .....	259
Kāpēc semantiskās tehnoloģijas būvniecības nozarē neattaisno cerības? .....	261
<b>NODAĻA 7.2. SLĒGTO PROJEKTU FORMĀTI UN SADARBSPĒJAS JAUTĀJUMI .....</b>	<b>265</b>
Slēgti dati un samazināta produktivitāte: CAD nozares strupceļš (BIM).....	265
Mīts par sadarbspēju starp CAD sistēmām .....	267
Pārejiet uz USD un granulāriem datiem m.....	271

NODAĻA 7.3. GEOMETRIJA BŪVNIECĪBĀ: NO LĪNIJĀM LĪDZ KUBIKMETRIEM .....	275
Kad līnijas pārvēršas naudā jeb kāpēc celtniekiem nepieciešama ģeometrija .....	275
No līnijām uz apjomiem: kā platība un apjoms kļūst par datiem .....	275
Pāreja uz MESH, USD un daudzstūriem: ģeometrijai izmantojiet teselāciju .....	277
LOD, LOI, LOMD - unikāla datu klasifikācija CAD (BIM). ....	279
Jaunie CAD standarti (BIM) - AIA, BEP, IDS, LOD, COBie.....	282
NODAĻA 7.4. PROJEKTĒŠANAS PARAMETRU NOTEIKŠANA UN LLM IZMANTOŠANA KAD DARBĪBAI	287
CAD datu unikalitātes ilūzija (BIM): ceļš uz analītiku un atvērtiem formātiem .....	287
Projektēšana, izmantojot parametrus: CAD nākotne un BIM.....	290
LLM parādīšanās projektēšanas CAD datu apstrādes procesos.....	293
Automatizēta DWG failu analīze ar LLM un Pandas .....	296
Nākamie soļi: pāreja no slēgtiem formātiem uz atvērtiem datiem .....	302
<b>VIII DAĻA UZ DATIEM BALSTĪTA LĒMUMU PIENEMŠANA, ANALĪTIKA, AUTOMATIZĀCIJA UN MAŠINMĀCĪŠANĀS.....</b>	<b>304</b>
NODAĻA 8.1. DATU ANALĪTIKA UN UZ DATIEM BALSTĪTU LĒMUMU PIENEMŠANA. ....	305
Dati kā resursu lēmumu pieņemšanā .....	305
datu vizualizācija: izpratnes un lēmumu pieņemšanas atslēga .....	308
KPI un ROI.....	310
Informācijas paneli un informācijas paneli: metriku vizualizēšana efektīvai pārvaldībai.....	312
Datu analīze un māksla uzdot jautājumus .....	314
NODAĻA 8.2. DATU PLŪSMA BEZ MANUĀLAS PIEPŪLES: KĀPĒC NEPIECIEŠAMS ETL.....	317
ETL automatizācija: izmaksu samazināšana un datu apstrādes paātrināšana.....	317
ETL Izraksts: datu vākšana.....	321
ETL Transform: validācijas un transformācijas noteikumu piemērošana. ....	324
ETL Load: Vizualizējet rezultātus diagrammās un grafikos. ....	326
ETL Load: Automātiska PDF dokumentu izveide.....	331
ETL Load: automātiska dokumentu ģenerēšana no FPDF.....	332
ETL Load: Ziņošana un ielāde uz citām sistēmām. ....	336
ETL ar LLM: datu vizualizēšana no PDF -dokumentiem .....	337
NODAĻA 8.3. AUTOMĀTISKAIS ETL KONVEIJERS (CAURUĻVADS) .....	342
Cauruļvads: Automātiskais ETL konveijers dati.....	342
Cauruļvads -ETL datu validācijas process ar LLM .....	346
Cauruļvads -ETL: projekta elementu datu un informācijas pārbaude CAD (BIM).....	348
NODAĻA 8.4. ETL UN DARBA PLŪSMU ORKESTRĒŠANA: PRAKTISKI RISINĀJUMI.....	354

DAG un Apache Airflow: darba plūsmas automatizācija un orkestrācija.....	354
Apache Airflow: praktisks ETL automatizācijā.....	355
Apache NiFi maršrutēšanai un datu konvertēšanai .....	359
n8n Low-Code, No-Code procesu orkestrācija.....	360
Nākamie soji: pāreja no manuālām operācijām uz analītikā balstītiem risinājumiem .....	363
<b>IX DAĻA DATU GLABĀŠANA UN PĀRVALDĪBA BŪVΝIECĪBĀ.....</b>	<b>365</b>
<b>NODAĻA 9.1. DATU INFRASTRUKTŪRA: NO GLABĀŠANAS FORMĀTIEM LĪDZ DIGITĀLAJĀM KRĀTUVĒM.....</b>	<b>366</b>
Datu atomi: efektīvas informācijas pārvaldības pamats .....	366
Informācijas glabāšana: faili vai dati .....	367
Lielo datu glabāšana: populāru formātu un to efektivitātes analīze .....	369
Uzglabāšanas optimizēšana ar Apache Parquet .....	372
DWH: Datu noliktava datu noliktavas .....	374
Datu ezers - no ETL uz ELT: no tradicionālās tīrišanas uz elastīgu apstrādi .....	376
Datu ezernīcas arhitektūra: noliktavu un datu ezeru sinerģija .....	377
CDE, PMIS, ERP vai DWH un datu ezers .....	380
<b>NODAĻA 9.2. DATU NOLIKTAVAS PĀRVALDĪBA UN HAOSA NOVĒRŠANA .....</b>	<b>383</b>
Vektoru datubāzes un robežlaukums.....	383
Datu pārvaldība), datu minimālisms) un datu purvs). .....	386
DataOps un VectorOps: jaunie datu standarti .....	389
Nākamie soji: no haotiskas glabāšanas uz strukturētu glabāšanu .....	391
<b>X DAĻA LIELIE DATI, MAŠINMĀCIŠANĀS UN PROGNOZES .....</b>	<b>393</b>
<b>NODAĻA 10.1. LIELIE DATI UN TO ANALĪZE .....</b>	<b>394</b>
Lielie dati būvniecībā: no intuīcijas līdz paredzamībai .....	394
Jautājums par lielo datu izmantošanas iespējamību: korelācija, statistika un datu izlase .....	395
Lielie dati: datu analīze no Sanfrancisko miljoniem būvatļauju datu kopas .....	398
Lielo datu piemērs, pamatojoties uz CAD datiem (BIM).....	404
IoT Lietu internets un viedie līgumi.....	408
<b>NODAĻA 10.2. MAŠINMĀCIŠANĀS UN PROGNOZES .....</b>	<b>412</b>
Mašinmācišanās un mākslīgais intelekts mainīs veidu, kā mēs veidojam. .....	412
No subjektīva novērtējuma līdz statistiskai prognozei.....	414
Titānika datu kopa: Hello World analītikas pasaule dati un lielie dati .....	416
Mašinmācišanās darbībā: no "Titānika" pasažieriem līdz projektu pārvaldībai .....	421
Prognozes un prognozes, pamatojoties uz vēsturiskiem datiem.....	425

Mašīnmācīšanās pamatjēdzeni .....	427
<b>NODAĻA 10.3. IZMAKSU UN TERMIŅU PROGNOZĒŠANA, IZMANTOJOT MAŠĪNMĀCĪŠANOS.....</b>	<b>430</b>
Piemērs mašīnmācīšanās izmantošanai, lai atrastu projekta izmaksas un termiņus. ....	430
Projekta izmaksu un laika prognozēšana, izmantojot lineāro regresiju .....	432
Projekta izmaksu un laika prognozes, izmantojot K-tuvāko kaimiņu algoritmu (k-NN). ....	435
Nākamie soļi: no uzglabāšanas līdz analīzei un prognozēšanai .....	439
<b>XI DAĻA BŪVΝIECĪBAS NOZARE DIGITĀLAJĀ LAIKMETĀ. IESPĒJAS UN IZAICINĀJUMI.....</b>	<b>442</b>
<b>NODAĻA 11.1. IZDŽĪVOŠANAS STRATĒGIJAS: KONKURENCES PRIEKŠROCĪBU VEIDOŠANA.....</b>	<b>443</b>
Korelācijas aprēķinu vietā: būvniecības analītikas nākotne.....	443
Uz datiem balstīta pieeja būvniecībā: jauns infrastruktūras līmenis .....	446
Nākamās paaudzes digitālais birojs: kā mākslīgais intelekts maina darba vidi.....	448
Atvērtie dati un Uberizācija ir drauds esošajiem būvniecības uzņēmumiem .....	450
Neatrisinātās uberizācijas problēmas kā pēdējā iespēja izmantot laiku pārveidošanai .....	453
<b>NODAĻA 11.2. PRAKTISKS CEĻVEDIS UZ DATIEM BALSTĪTAS PIEEJAS IEVIEŠANAI .....</b>	<b>458</b>
No teorijas uz praksi: ceļvedis digitālajai transformācijai būvniecībā .....	458
Digitālā pamata veidošana: 1-5 soļi ceļā uz digitālo briedumu.....	460
Datu potenciāla atraisīšana: 5-10 soļi uz digitālo briedumu.....	464
Transformācijas ceļvedis: no haosa līdz uz datiem balstītam uzņēmumam .....	471
Būvniecība 5.0 industrijā: kā pelnīt naudu, kad vairs nav iespējams slēpties .....	474
<b>SECINĀJUMS .....</b>	<b>476</b>
<b>PAR AUTORU.....</b>	<b>479</b>
<b>TURPMĀKĀS ATTIECĪBAS .....</b>	<b>480</b>
<b>TULKOJUMA KOMENTĀRS .....</b>	<b>480</b>
<b>CITAS PRASMES UN JĒDZIENI .....</b>	<b>481</b>
<b>GLOSSĀRIJS .....</b>	<b>485</b>
<b>ATSAUČU UN TIEŠSAISTES MATERIĀLU SARAKSTS .....</b>	<b>491</b>
<b>SUBJEKTU INDEKSS.....</b>	<b>509</b>



## MAKSIMĀLI ĒRTA DRUKĀTĀ VERSIJA

Jūsu rīcībā ir bezmaksas digitālā versija **Data-Driven Construction**. Ērtākam darbam un ātrākai piekļuvei materiāliem iesakām pievērst uzmanību **drukātajam izdevumam**:



■ **Vienmēr pa rokai:** grāmata drukātā formātā klūs par uzticamu darba rīku, kas ļaus ātri atrast un izmantot vajadzīgās vizualizācijas un diagrammas jebkurā darba situācijā.

■ **Augstas kvalitātes ilustrācijas:** visi attēli un grafikas drukātajā izdevumā ir maksimāli kvalitatīvi.

■ **Ātra piekļuve informācijai:** vienkārša navigācija, iespēja veikt piezīmes, grāmatzīmes un strādāt ar grāmatu jebkurā vietā.

Iegādājoties pilnu drukāto grāmatas versiju, jūs iegūstat ērtu rīku ērtam un efektīvam darbam ar informāciju: iespēju ātri izmantot vizuālos materiālus ikdienas uzdevumos, ātri atrast nepieciešamās shēmas un veikt piezīmes. Turklāt ar savu pirkumu atbalstāt atvērto zināšanu izplatīšanu.

Grāmatas drukāto versiju var pasūtīt: [datadrivenconstruction.io/books](http://datadrivenconstruction.io/books)



## I DAĻA

# **NO MĀLA PLĀKSNĒM LĪDZ DIGITĀLAJAI REVOLŪCIJAI: KĀ INFORMĀCIJA IR ATTĪSTĪJUSIES BŪVNIECĪBĀ**

Grāmatas pirmajā daļā aplūkota datu pārvaldības vēsturiskā attīstība būvniecības nozarē, sākot no primitīviem ierakstiem fiziskos datu nesējos līdz mūsdienu digitālajām ekosistēmām. Tajā analizēta informācijas pārvaldības tehnoloģiju transformācija, ERP -sistēmu rašanās un datu fragmentācijas ietekme uz uzņēmējdarbības procesu efektivitāti. Īpaša uzmanība tiek pievērsta informācijas digitalizācijas procesam un objektīvas analīzes pieaugošajai nozīmei, aizstājot subjektīvu ekspertu vērtējumu. Detalizēti tiek aplūkots eksponenciālais informācijas apjoma pieaugums, ar ko saskaras mūsdienu būvniecības nozare, un ar to saistītās problēmas uzņēmumu sistēmām. Tieks pētīta būvniecības nozares pozicionēšana ceturtās un piektās industriālās revolūcijas kontekstā, kā arī mākslīgā intelekta un uz datiem orientētu pieeju izmantošanas potenciāls, lai radītu ilgtspējīgas konkurences priekšrocības.

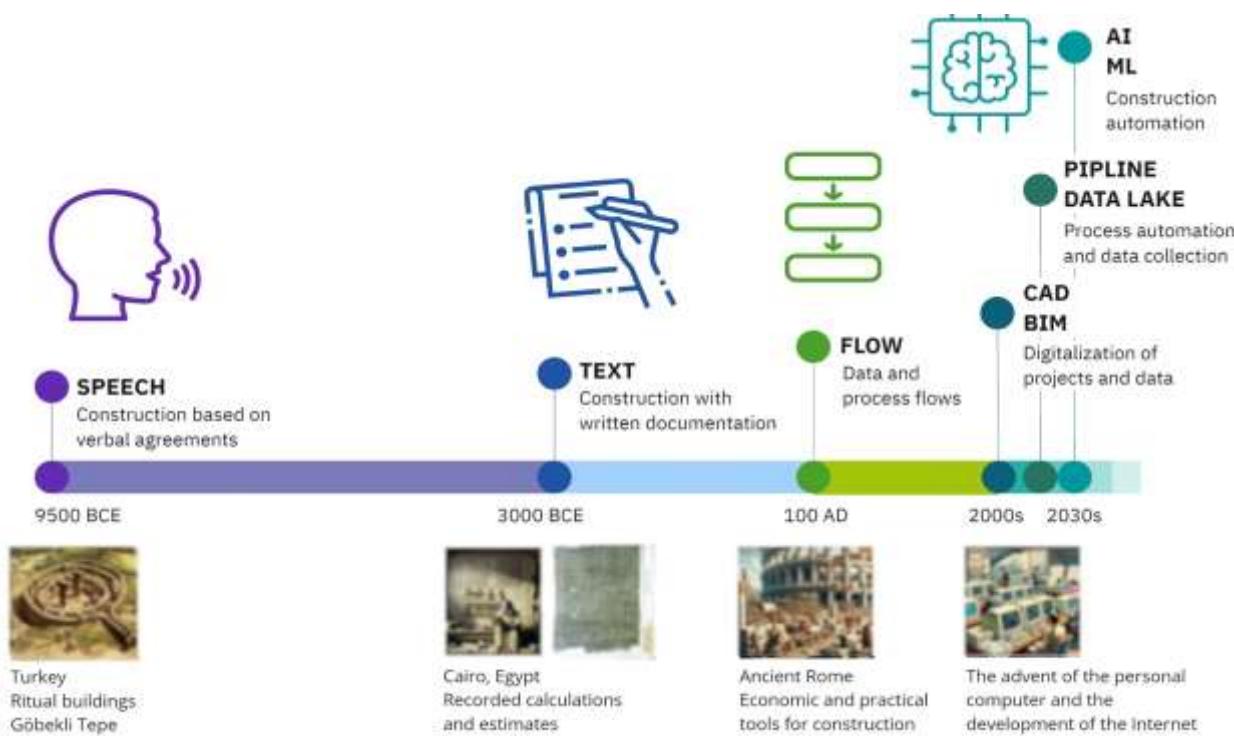
## NODĀLA 1.1.

### DATU IZMANTOŠANAS ATTĪSTĪBA BŪVΝIECĪBAS NOZARĒ

#### Datu ēras sākums būvniecībā

Pirms aptuveni 10 000 gadu, neolīta laikmetā, cilvēce veica revolucionāru pāreju savā attīstībā, atsakoties no nomadu dzīvesveida un pārejot uz mazkustīgu dzīvi, kā rezultātā parādījās pirmās primitīvās ēkas no māla, koka un akmens [6]. No šī brīža sākas būvniecības nozares vēsture.

Attīstoties civilizācijām, arhitektūra kļuva arvien sarežģītāka, un tā radās pirmie rituālie tempļi un sabiedriskās ēkas. Arhitektūras projektu pieaugošā sarežģītība lika antīkajiem inženieriem un vadītājiem izveidot pirmos pierakstus un aprēķinus. Pirmie pieraksti uz māla plāksnītēm un papirusiem bieži vien ietvēra loģikas aprakstu, aprēķinot nepieciešamo būvmateriālu daudzumu, to izmaksas un aprēķinot samaksu par paveikto darbu [7]. Tā sākās datu izmantošanas ēra būvniecībā - ilgi pirms mūsdienu digitālo tehnoloģiju parādīšanās (1.1.1-1. attēls).

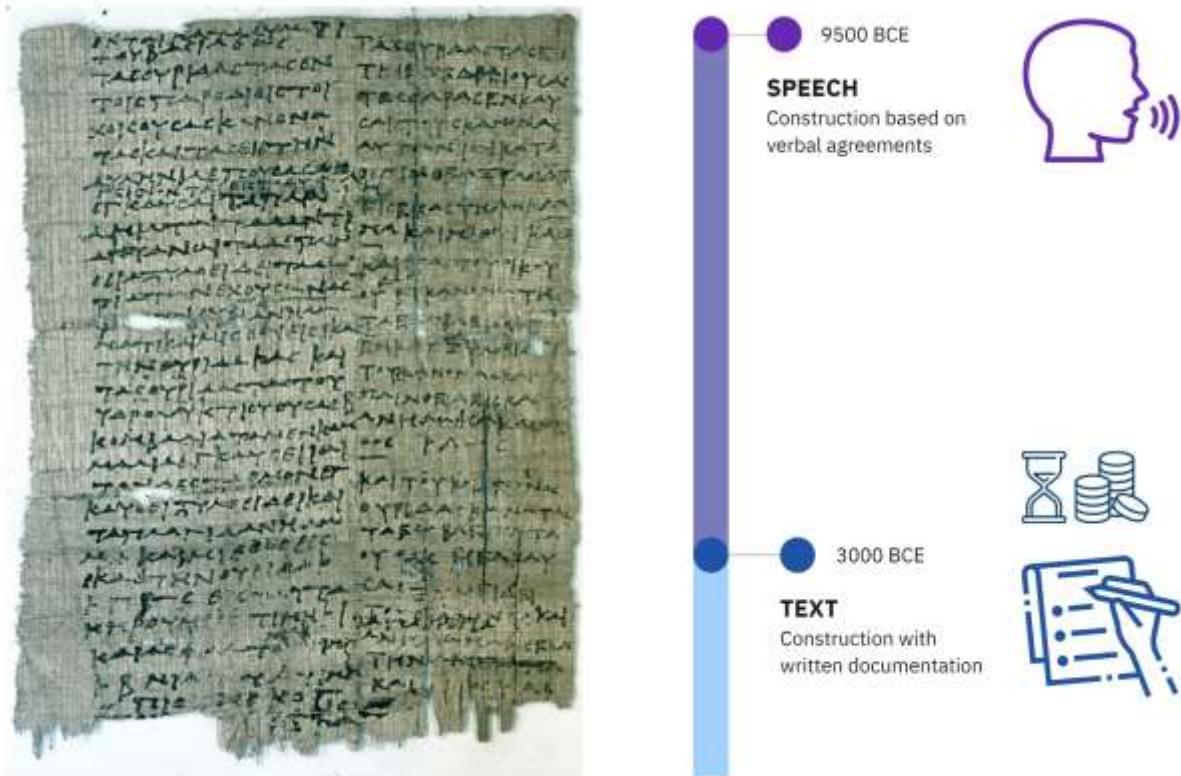


Attēls 1.1-1 Informācijas tehnoloģiju attīstības kronoloģija būvniecībā: no verbālās informācijas līdz mākslīgajam intelektam.

## No māla un papirusa līdz digitālajām tehnoloģijām

Pirmās dokumentārās liecības par būvniecību datējamas ar piramīdu celtniecības periodu, aptuveni 3000-4000 gadu pirms mūsu ēras[7]. Kopš tā laika rakstveida liecību glabāšana ir veicinājusi un pavadījusi progresu būvniecības nozarē, laujot uzkrāt un sistematizēt zināšanas, kas turpmāko 10 000 gadu laikā radīja būtiskus jauninājumus būvniecības metodēs un arhitektūrā.

Pirma fizisko datu nesēju izmantošana būvniecībā, piemēram, māla plāksnītes, papiruss pirms tūkstošiem gadu ( 1.1-2. attēls) vai "A0" papīrs 20. gadsimta 80. gados, lai reģistrētu datus, sākotnēji nebija paredzēta šīs informācijas izmantošanai jaunos projektos. Šādu ierakstu galvenais mērķis bija detalizēti atspoguļot projekta pašreizējo stāvokli, tostarp aprēķinus par nepieciešamajiem materiāliem un darbu izmaksām. Līdzīgi arī mūsdienās digitālo projekta datu un modeļu pieejamība ne vienmēr garantē to izmantošanu turpmākajos projektos un bieži vien kalpo galvenokārt kā informācija aktuāliem aprēķiniem par nepieciešamajiem materiāliem un būvniecības izmaksām.



Attēls 1.1-2 3. gadsimta p. m. ē. papiruss, kurā aprakstītas izmaksas, kas saistītas ar dažādu veidu logu apgleznošanu karāja pilī, izmantojot enkaustikas tehniku.

Cilvēcei vajadzēja aptuveni 5000 gadu, lai būvniecības projektu vadībā no mutiskām sarunām pāriņtu uz rakstiskiem dokumentiem, un tikpat ilgs laiks bija nepieciešams, lai pāriņtu no papīra uz digitāliem datiem kā galveno plānošanas un kontroles resursu.

Tāpat kā tirdzniecības un monetāro attiecību attīstība veicināja rakstības un pirmo juristu rašanos, lai risinātu strīdus, tā arī pirmie materiālo izmaksu un darbu apjomu ieraksti būvniecībā veicināja pirmo vadītāju rašanos būvniecības nozarē, kuru pienākumos ietilpa dokumentēšana, uzraudzība un atbildība par galveno informāciju par projekta termiņiem un izmaksām.

Mūsdienās datiem ir daudz nozīmīgāka loma: tie ne tikai reģistrē pieņemtos lēmumus, bet arī klūst par nākotnes prognozēšanas un modelēšanas rīku. Tas ir pamats, uz kura tiek veidota mūsdienīga procesu pieeja projektu vadībā, pārvēršot uzkrāto pieredzi lēmumu pieņemšanas sistēmā, kas balstīta uz strukturētiem un pārbaudāmiem datiem.

## **Process kā rīks uz datiem balstītai pieredzei**

Jebkura procesa pamatā ir pagātnes pieredzes pārveidošana par instrumentu nākotnes plānošanai. Pieredze mūsdienām izpratnē ir strukturēts datu kopums, ko var analizēt, lai veiktu pamatotas prognozes.

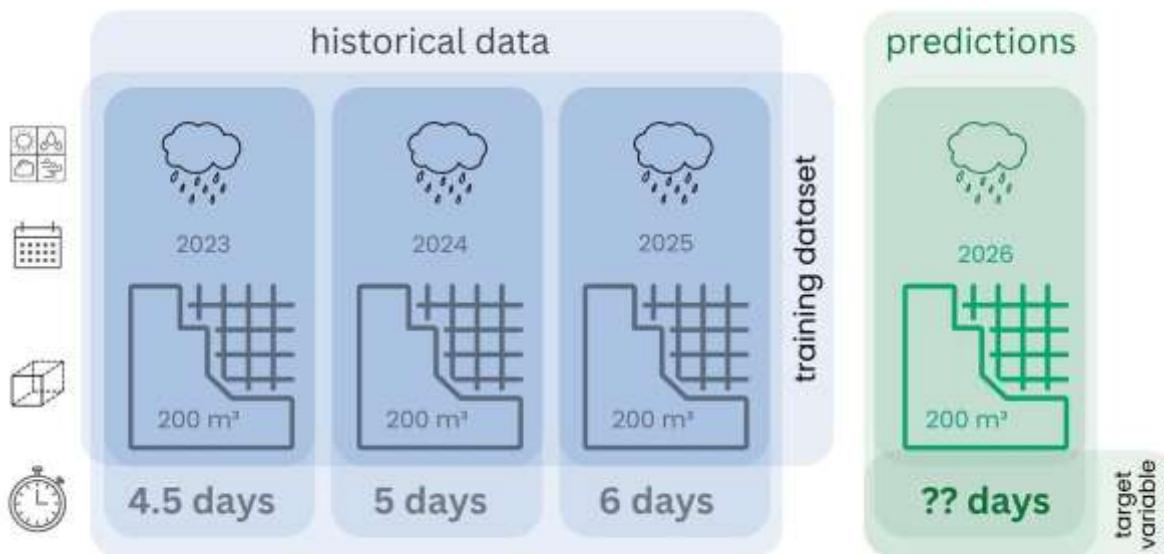
Tieši vēsturiskie dati ir prognožu pamatā, jo tie skaidri parāda veiktā darba rezultātus un sniedz ieskatu par faktoriem, kas ietekmē šos rezultātus.

Nemsim konkrētu piemēru no monolītās būvniecības: parasti, plānojot darbu grafiku, tiek ņemts vērā betona apjoms, konstrukcijas sarežģītība un laika apstākļi. Pieņemsim, ka konkrētais būvdarbu vadītājs vai uzņēmuma vēsturiskie dati par pēdējiem trim gadiem (2023-2025) liecina, ka 200 m<sup>2</sup> lielas monolītās konstrukcijas betonēšana lietainā laikā ilga no 4,5 līdz 6 dienām ( 1.1-3. attēls). Tieši šie uzkrātie statistikas dati klūst par pamatu, lai, plānojot līdzīgus darbus turpmākajos projektos, prognozētu izpildes laiku un aprēķinātu resursu izmaksas. Pamatojoties uz šiem vēsturiskajiem datiem, būvuzraugs vai tāmētājs, balstoties uz pieredzi, var veikt pamatotu prognozi par laiku, kas nepieciešams, lai 2026. gadā līdzīgos apstākļos pabeigtu līdzīgus darbus nākotnē.

Šajā laika analītikas gadījumā analītiskais process darbojas kā mehānisms, lai pārveidotu atšķirīgus datus strukturētā pieredzē un pēc tam precīzā plānošanas rīkā. Datu un procesi ir vienota ekosistēma, kur viens bez otra nevar pastāvēt.

*Saskaitiet to, kas ir saskaitāms, izmēriet to, kas ir izmērāms, un padariet izmērāmu to, kas nav izmērāms [8].*

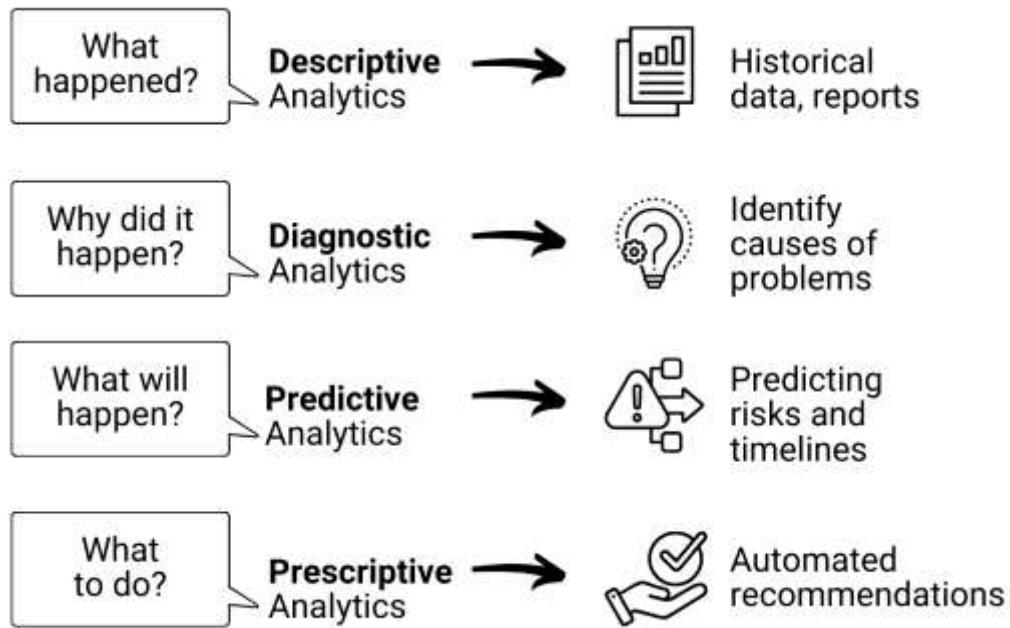
- Galileo Galilejs



Attēls 1.1-3 Vēsturiskie dati kalpo kā mācību datu kopa, lai prognozētu kādu no vērtībām nākotnē.

Mūsdienu uzņēmējdarbības vidē datu analīze kļūst par būtisku komponentu efektīvai projektu vadībai, procesu optimizācijai un stratēģisko lēmumu pieņemšanai. Būvniecības nozarē pakāpeniski tiek apgūti četri galvenie analītikas līmeni, no kuriem katrs atbild uz konkrētu jautājumu un sniedz unikālus ieguvumus ( 1.1-4. attēls):

- **Aprakstošā analīze** - sniedz atbildes uz jautājumu "kas notika?" un sniedz vēsturiskus datus un pārskatus par pagātnes notikumiem un rezultātiem: pēdējo trīs gadu laikā (2023-2025) lietus laikā 200m<sup>2</sup> monolītās konstrukcijas ielejšana aizņēma no 4,5 līdz 6 dienām.
- **Diagnostikas analītika** - atbild uz jautājumu "kāpēc tā notika?", nosakot problēmu cēloņus: analīze liecina, ka monolītās konstrukcijas lietainās ielejas laiks palielinājās lietus laika apstākļu dēļ, kas palēnināja betona sacietēšanas procesu.
- **Prognozēšanas analītika** - uz nākotni vērsta, paredzot iespējamos riskus un izpildes termiņus, atbildot uz jautājumu "kas notiks?": pamatojoties uz vēsturiskajiem datiem, tiek prognozēts, ka līdzīgas 200 m<sup>2</sup> lielas monolītas konstrukcijas betonēšana lietainā laikā 2026. gadā prasīs aptuveni 5,5 dienas, ņemot vērā visus zināmos faktorus un tendences.
- **Preskriptīvā analīze** - sniedz automatizētus ieteikumus un atbild uz jautājumu "ko darīt?", ņaujot uzņēmumiem izvēlēties optimālo rīcību: Lai optimizētu darbus, piemēram, ieteicams: izmantot īpašas piedevas, lai paātrinātu betona sacietēšanu augsta mitruma apstākjos; plānot betonēšanu periodiem ar vismazāko nokrišņu varbūtību; ierīkot konstrukcijai pagaidu nojumes, kas pat nelabvēlīgos laika apstākļos samazinās darbu laiku līdz 4-4,5 dienām.



Attēls 1.1-4 Galvenie analītikas veidi: no pagātnes apraksta līdz automatizētai lēmumu pieņemšanai.

Pilnvērtīgai digitālajai transformācijai, kas nozīmē pāreju uz sistēmas analītiku un uz datiem balstītu pārvaldību, ir nepieciešama ne tikai ārpakalpojumu izmantošana, bet arī kompetentas iekšējās komandas izveide. Šādas komandas galvenajiem locekļiem jābūt produktu vadītājiem, datu inženieriem, analītiķiem un izstrādātājiem, kuri strādās ciešā sadarbībā ar uzņēmējdarbības struktūrvienībām ( 4.3-9. attēls). Šāda sadarbība ir nepieciešama, lai uzdotu gudrus analītiskus jautājumus un efektīvi parametrizētu biznesa lēmumu pieņemšanas uzdevumus. Informācijas sabiedrībā dati kļūst ne tikai par palīginstrumentu, bet gan par prognozēšanas un optimizācijas pamatu.

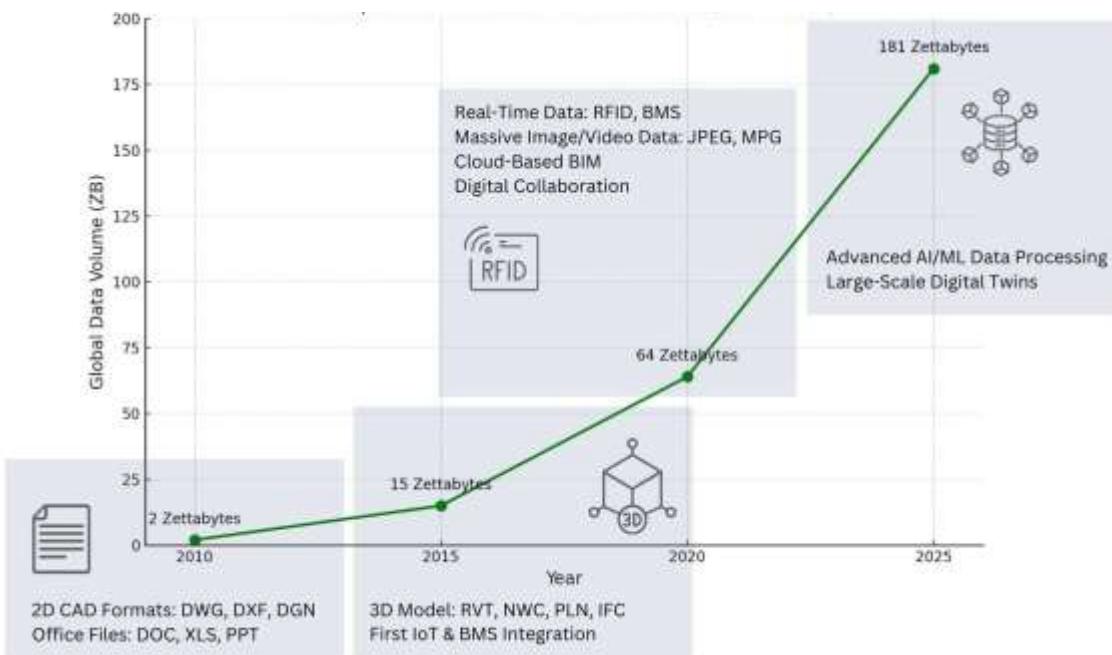
Būvniecībā digitālā transformācija būtiski maina veidu, kā tiek projektēti, pārvaldīti un ekspluatēti objekti. Šo procesu dēvē par informācijas digitalizāciju, kad visi būvniecības procesa aspekti tiek pārvērsti digitālā formā, kas piemērota analīzei.

## Būvniecības procesa informācijas digitalizācija

Gadsimtiem ilgi būvniecībā reģistrētās informācijas apjoms ir saglabājies gandrīz nemainīgs, taču pēdējās desmitgadēs tas ir strauji pieaudzis ( 1.1.1-5. attēls).

Saskaņā ar PwC pētījumu® "Pārvaldītie dati. Kas studentiem nepieciešams, lai gūtu panākumus strauji mainīgajā biznesa pasaule" (2015) [9], 90% no visiem datiem pasaule ir radīti pēdējo divu gadu laikā (2015. gadā). Tomēr lielākā daļa uzņēmumu šos datus pilnībā neizmanto, jo tie vai nu paliek atsevišķas sistēmās, vai tiek vienkārši arhivēti bez reālas analīzes.

Datu apjoma pieaugums pēdējos gados ir tikai paātrinājies - no 15 zetabaitiem 2015. gadā līdz 181 zetabaitam 2025. gadā [10]. Būvniecības un projektēšanas uzņēmumu serveri katru dienu tiek piepildīti ar projekta dokumentāciju, darbu grafikiem, aprēķiniem un aprēķiniem, finanšu pārskatiem. 2D/3D - rasējumiem izmanto DWG, DXF un DGN formātus, bet 3D modeliem - RVT, NWC, PLN un IFC™. Teksta dokumenti, tabulas un prezentācijas tiek saglabāti DOC, XLSX un PPT formātā. Papildus video un attēliem no būvlaukuma - MPG un JPEG formātā, reāllaika dati no IoT komponentiem, RFID® tagiem (identifikācija un izsekošana) un BMS ēku vadības sistēmām (uzraudzība un kontrole)



Attēls 1.1.1-5 Paraboliskais datu pieaugums 2010-2025 (pamatojoties uz [10]).

Strauji pieaugot informācijas apjomam, būvniecības nozare saskaras ar nepieciešamību ne tikai vākt un uzglabāt datus, bet arī nodrošināt to pārbaudi, validāciju, izmērāmību un analītisku apstrādi. Pašlaik nozarē notiek aktīvs informācijas digitalizācijas posms - sistemātiska visu būvniecības darbības aspektu pārveidošana digitālā formā, kas piemērota analīzei, interpretācijai un automatizācijai.

Informācijas digitalizācija nozīmē ņemt informāciju par visām būvniecības projekta vienībām un elementiem un pašu būvniecības procesu - arī par tiem, kurus iepriekš vispār neuzskatījām par informāciju, - un pārvērst to datu formātā, lai informācija būtu kvantitatīvi izmērāma un viegli analizējama.

Būvniecības kontekstā tas nozīmē apkopot un digitalizēt informāciju visiem projektu elementiem un

visiem procesiem - sākot no tehnikas un cilvēku kustības būvlaukumā līdz laikapstākļiem un klimatis-kaijiem apstākļiem būvlaukumā, pašreizējām materiālu cenām un centrālo banku procentu likmēm -, lai radītu analītiskus modeļus.

*Ja jūs varat izmērīt to, par ko jūs runājat, un izteikt to skaitlos, tad jūs kaut ko zināt par šo tēmu. Bet, ja jūs to nevarat izteikt kvantitatīvi, jūsu zināšanas ir ļoti ierobežotas un neapmierinošas. Tas var būt sākumpunkts, bet tas nav īsto zinātnisko zināšanu līmenis. [11].*

- V. Tomsons (lords Kelvins), 1824-1907, britu zinātnieks.

Informācijas digitalizācija ir daudz plašāka par tradicionālo pieeju informācijas vākšanai, kad tika reģistrēti tikai tādi pamatrādītāji kā cilvēkstundas vai faktiskās materiālu izmaksas. Mūsdienās praktiski jebkuru notikumu var pārvērst datu plūsmā, kas piemērota padzilinātai analīzei, izmantojot progresīvus analīzes rīkus un mašīnmācīšanās metodes. Būvniecības nozarē ir notikusi fundamentāla pāreja no papīra rasējumiem, Excel tabulām un mutiskiem norādījumiem uz digitālām sistēmām ( 1.2-4. attēls), kurās katrs projekta elements klūst par datu avotu. Pat darbinieki - no inženieriem līdz būvdarbu strādniekiem būvlaukumā - tagad tiek uzskatīti par digitālu mainīgo lielumu un datu kopu kopumu.

Saskaņā ar KPMG pētījuma "Familiar Challenges - New Approaches: Global Construction Survey 2023" ("Pazīstami izaicinājumi - jaunas pieejas: pasaules būvniecības apsekojums 2023") datiem digitālie dvīņi, mākslīgais intelekts (AI) un lielie dati, klūst par galvenajiem virzītājspēkiem projektu rentabilitātes uzlabošanai [2].

Modernās tehnoloģijas ne tikai vienkāršo informācijas vākšanu, padarot to lielā mērā automātisku, bet arī radikāli samazina datu uzglabāšanas izmaksas. Tā rezultātā uzņēmumi atsakās no selektīvas pieejas un dod priekšroku visa informācijas kopuma uzglabāšanai, lai vēlāk to varētu analizēt ( 2.1.1-5. attēls), kas paver potenciālas iespējas optimizēt procesus nākotnē.

Informācijas digitalizācija un digitalizācija ļauj atklāt apslēpto, iepriekš neizmantoto informācijas vērtību. Ja dati ir pareizi organizēti, tos var atkārtoti izmantot, atkārtoti interpretēt un integrēt jaunos pakalpojumos un risinājumos.

Nākotnē informācijas digitalizācija, visticamāk, novedīs pie pilnīgas dokumentu pārvaldības automātizācijas, pašvadāmu būvniecības procesu ieviešanas un jaunu profesiju - būvniecības datu analītiķu, mākslīgā intelekta projektu vadības ekspertu un digitālo inženieru - rašanās. Būvniecības projekti klūs par dinamiskiem informācijas avotiem, un lēmumu pieņemšana balstīsies nevis uz intuīciju vai subjektīvu pieredzi, bet gan uz ticamiem un reproducējamiem digitāliem faktiem

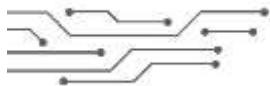
*Informācija ir 21. gadsimta nafta, un analītika ir iekšdedzes dzinējs [12].*

- Peter Sondergaard, Gartner vecākais viceprezidents.®

Saskaņā ar IoT Analytics 2024 [13] datiem paredzams, ka globālie izdevumi datu pārvaldībai un analīzei krasī pieauga no 185,5 miljardiem ASV dolāru 2023. gadā līdz 513,3 miljardiem ASV dolāru 2030. gadā, sasniedzot 16% gada pieauguma līmeni. Tomēr ne visi komponenti aug vienādā tempā: analītikas strauji pieaug, savukārt datu glabāšanas pieaugums palēninās. Analītika nodrošinās visstraujāko datu pārvaldības ekosistēmas izaugsmi: tiek prognozēts, ka tā pieauga no 60,6 miljardiem ASV dolāru 2023. gadā līdz 227,9 miljardiem ASV dolāru 2030. gadā, kas ir 27% pieaugums gadā.

Līdz ar paātrināto informācijas digitalizāciju un straujo informācijas apjoma pieaugumu būvniecības projektu un uzņēmumu vadība saskaras ar nepieciešamību sistemātiski uzglabāt, analizēt un apstrādāt dažādus, bieži vien neviendabīgus datus. Reaģējot uz šo izaicinājumu, sākot ar 90. gadu vidu, nozarē sākās masveida pāreja uz elektronisku dokumentācijas izveidi, glabāšanu un pārvaldību - sākot no elektroniskajām tabulām un projekta aprēķiniem līdz rasējumiem un līgumiem.

Tradicionālos papīra dokumentus, kuriem nepieciešami paraksti, fiziska glabāšana, regulāra pārskatīšana un arhivēšana skapjos, pakāpeniski aizstāj digitālās sistēmas, kas datus glabā strukturētā veidā - specializētu lietojumprogrammu datubāzēs.



## NODAĻA 1.2.

### TEHNOLOGIJAS UN VADĪBAS SISTĒMAS MŪSDIENU BŪVNIECĪBĀ

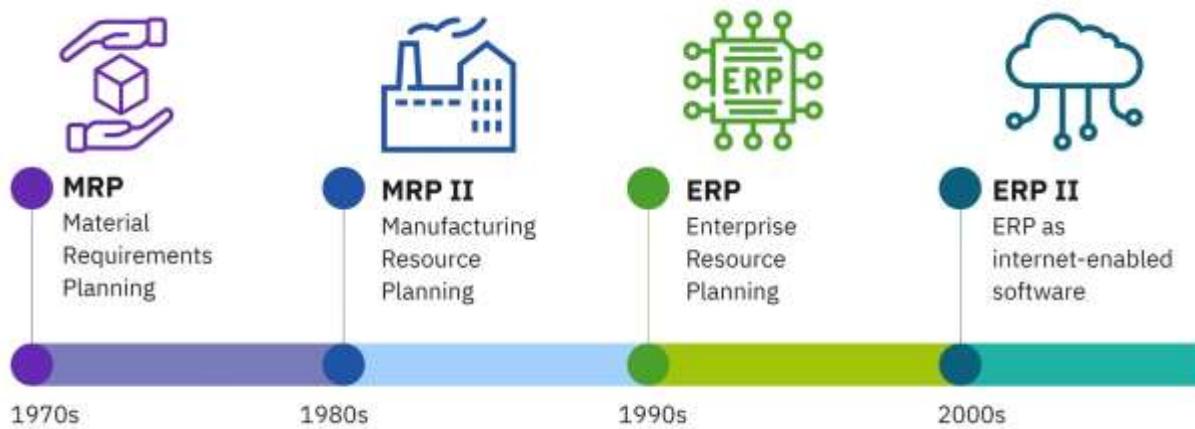
#### Digitālā revolūcija un moduļu MRP/ERP -sistēmu rašanās

Mūsdieno digitālo datu glabāšanas un apstrādes laikmets sākās ar magnētiskās lentes parādīšanos pagājušā gadsimta 50. gados, kas pavēra iespēju uzglabāt un izmantot lielus informācijas apjomus. Nākamais izrāviens bija disku diskdzīņu parādīšanās, kas radikāli mainīja pieejumu datu pārvaldībai būvniecības nozarē.

Attīstoties datu uzglabāšanai, daudzi uzņēmumi ir ienākuši risinājumu tirgū un sākuši izstrādāt modulāru programmatūru datu izveidei, uzglabāšanai, apstrādei un rutīnas uzdevumu automatizēšanai

Eksponenciālais informācijas un rīku pieaugums ir radījis nepieciešamību pēc integrētiem, modulāriem risinājumiem, kas nestrādā ar atsevišķiem failiem, bet palīdz pārvaldīt un kontrolēt datu plūsmu dažādos procesos un projektos.

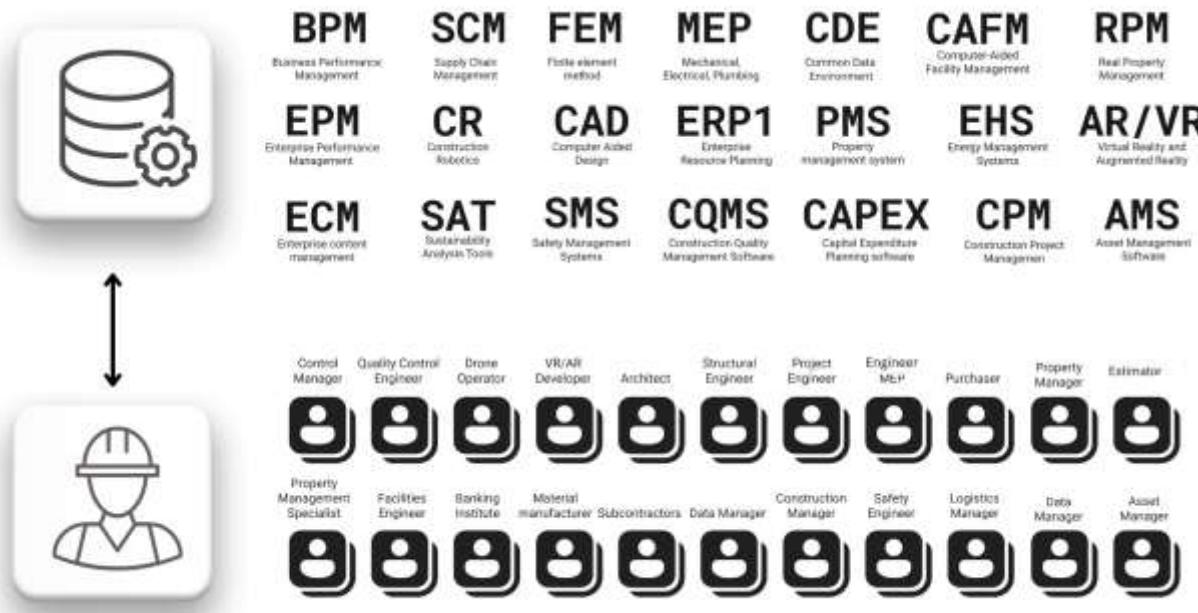
Pirmajiem visaptverošajiem platformas rīkiem bija ne tikai jāglabā dokumenti, bet arī jādokumentē visi izmaiņu pieprasījumi un darbības procesos: kas tos ierosināja, kāda bija pieprasījuma darbības joma un kas galu galā tika reģistrēts kā vērtība vai atribūts. Šiem nolūkiem bija nepieciešama sistēma, kas varētu izsekot precīziem aprēķiniem un pieņemtajiem lēnumiemiem (1.2-1. attēls). Šādas platformas bija pirmās MRP (Material Requirements Planning) un ERP (Enterprise Resource Planning) sistēmas, kas kļuva populāras no 90. gadu sākuma [14]



Attēls 1.2-1 Datu glabāšanas tehnoloģiju attīstība veicināja ERP -sistēmu rašanos 80. gados.

Pirmās MRP - un ERP - sistēmas lika pamatus digitalizācijas ērai biznesa procesu un būvniecības projektu vadībā. Modulārās sistēmas, kas sākotnēji bija paredzētas galveno uzņēmējdarbības procesu automatizēšanai, ar laiku tika integrētas ar papildu, elastīgākiem un pielāgojamākiem programmatūras risinājumiem.

Šie papildu risinājumi bija paredzēti datu apstrādei un projektu satura pārvaldībai ( 1.2-2. attēls), tie vai nu aizstāja atsevišķus lielo sistēmu moduļus, vai arī efektīvi papildināja tos, paplašinot visas sistēmas funkcionalitāti.



Attēlā 1.2-2. attēls Jaunie programmatūras risinājumi ir piesaistījuši uzņēmumā vadītāju armiju, lai pārvaldītu datu plūsmas.

Pēdējās desmitgadēs uzņēmumi ir veikuši lielus ieguldījumus moduļu sistēmās [15], uztverot tās kā ilgtermiņa integrētus risinājumus.

Saskaņā ar Software Path ziņojumu par 2022. gadu [16] vidējais budžets vienam ERP sistēmas lietotājam ir 9 000 ASV dolāru. Vidēji šādas sistēmas izmanto aptuveni 26% uzņēmuma darbinieku. Tādējādi organizācijai ar 100 lietotājiem kopējās ERP ieviešanas izmaksas sasniedz aptuveni 900 000 ASV dolāru.

Nemot vērā moderno, elastīgo un atvērto tehnoloģiju straujo attīstību, ieguldījumi patentētos, slēgtos un modulāros risinājumos kļūst arvien mazāk pamatoti. Ja šādi ieguldījumi jau ir veikti, ir svarīgi objektīvi pārvērtēt esošo sistēmu lomu: vai tās joprojām ir būtiskas ilgtermiņā, vai arī to funkcijas var pārskatīt un īstenot efektīvāk un pārredzamāk.

Viena no galvenajām mūsdieniņu modulāro datu platformu problēmām ir tā, ka tās centralizē datu pārvaldību slēgtās lietojumprogrammās. Tā rezultātā dati - uzņēmuma pamatvērtība - kļūst atkarīgi no konkrētiem programmatūras risinājumiem, nevis otrādi. Tas ierobežo informācijas atkalizmantošanu, apgrūtina migrāciju uz un samazina uzņēmējdarbības veiklību strauji mainīgajā digitālajā vidē.

Ja ir iespējams, ka nākotnē slēgtas modulārās arhitektūras vērtība vai nozīme mazināsies, ir lietderīgi atzīt, ka šodien radušās izmaksas ir neatgūstamas izmaksas, un koncentrēties uz stratēģisku pāreju uz atvērtāku, mērogojamu un pielāgojamu digitālo ekosistēmu.

Patentētu programmatūru raksturo izstrādātāja ekskluzīva kontrole pār pirmkodu un lietotāja datiem, kas radīti, izmantojot šādus risinājumus. Atšķirībā no atvērtā pirmkoda programmatūras lietotājiem nav pieķuves lietojumprogrammas iekšējai struktūrai, un viņi nevar patstāvīgi pārskatīt, pārveidot vai pielāgot to savām vajadzībām. Tā vietā viņiem ir jāiegādājas licences, kas dod tiesības lietot programmatūru ražotāja noteiktajās robežās.

Mūsdienīga uz datiem orientēta pieeja piedāvā citādu paradigmu: dati jāuzskata par galveno stratēģisko aktīvu - neatkarīgu, noturīgu un neatkarīgu no konkrētiem programmatūras risinājumiem. Savukārt lietojumprogrammas kļūst tikai par datu rīkiem, kurus var brīvi nomainīt, neriskējot zaudēt svarīgu informāciju.

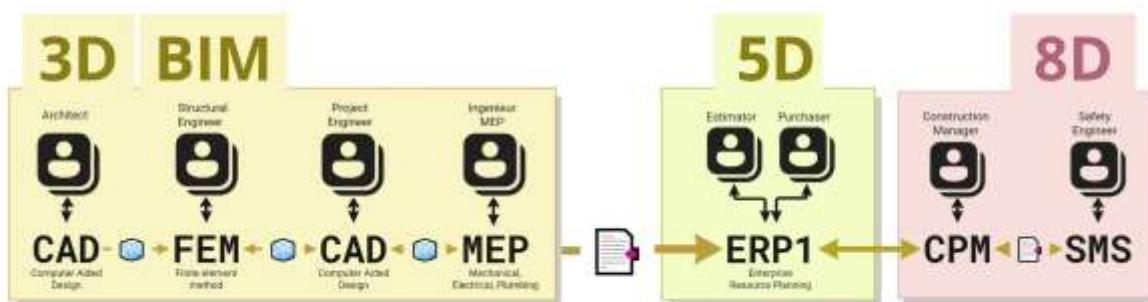
ERP un MRP sistēmu attīstība 90. gados (1.2-1. attēls) nodrošināja uzņēmumus ar jaudīgiem procesu pārvaldības rīkiem, taču tai bija arī neparedzētas sekas - ievērojami palielinājās informācijas plūsmu uzturēšanā iesaistīto cilvēku skaits. Tā vietā, lai automatizētu un vienkāršotu darbības uzdevumus, šīs sistēmas bieži vien radīja jaunu sarežģītības, birokrātijas un atkarības līmeni no iekšējiem IT resursiem.

### Datu pārvaldības sistēmas: no datu ieguves līdz uzņēmējdarbības izaicinājumiem

Mūsdieni uzņēmumi saskaras ar nepieciešamību integrēt vairākas datu pārvaldības sistēmas. Datu pārvaldības sistēmu izvēle, šo sistēmu pārdomāta pārvaldība un atšķirīgu datu avotu integrācija kļūst kritiski svarīga uzņēmējdarbības veikšanai.

20. gadsimta vidū var atrast simtiem (un lielos būvniecības uzņēmumos - tūkstošiem) dažādu sistēmu (1.2-3. attēls), kurām ir jāsadarbojas, lai visi būvniecības procesa aspekti noritētu raiti un saskaņoti.

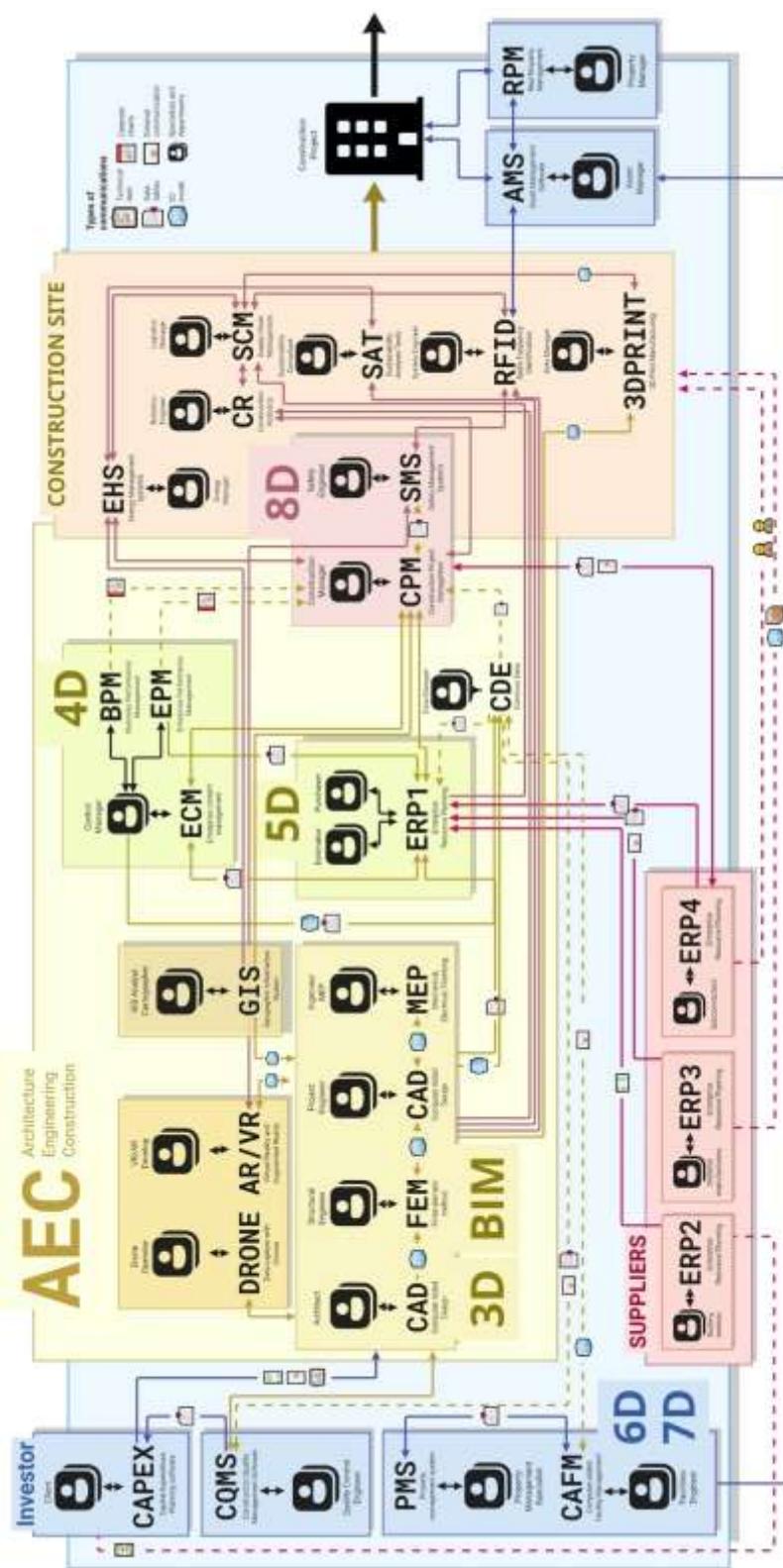
Saskaņā ar Deloitte 2016. gada pētījumu® "Data-Driven Management in Digital Capital Projects" (Datu pārvaldība digitālajos kapitālprojektos) vidējais būvniecības speciālists ikdienā izmanto 3,3 programmatūras lietojumprogrammas, bet tikai 1,7 no tām ir savstarpēji integrētas [17].



Attēls 1.2-3 Katrai biznesa sistēmai nepieciešama profesionāla komanda un atbildīgs vadītājs, kas nodrošina kvalitatīvu datu pārvaldību.

Turpmāk ir uzskaitītas populāras sistēmas, kas tiek izmantotas būvniecības nozares vidējiem un lieliem uzņēmumiem un ko izmanto efektīvai būvniecības projektu vadībai:

- **ERP (Uzņēmumu resursu plānošana)** - nodrošina biznesa procesu integrāciju, tostarp grāmatvedības, iepirkumu un projektu vadības procesus.
- **CAPEX (Kapitāla izdevumu plānošanas programmatūra)** - tiek izmantota, lai veidotu budžetu un pārvaldītu finanšu ieguldījumus būvniecības projektos, palīdz noteikt pamatlīdzekļu izmaksas un ieguldījumus ilgtermiņa aktīvos.
- **CAD (Computer-Aided Design) un BIM (Building Information Modeling)** - tiek izmantoti, lai izveidotu detalizētus un precīzus tehniskos rasējumus un 3D - projektu modeļus. Šīs sistēmas koncentrējas uz darbu ar ģeometrisko informāciju.
- **MEP (mehāniskā, elektriskā un santehnikas)** - inženiertehniskās sistēmas, kas ietver mehāniskās, elektriskās un santehnikas sastāvdajās un detalizēti apraksta projekta iekšējo "asinsrites" sistēmu.
- **GIS (ģeogrāfiskās informācijas sistēmas)** - izmanto reljefa analīzei un plānošanai, tostarp kartogrāfijai un telpiskajai analīzei.
- **CQMS (būvniecības kvalitātes vadības programmatūra)** - nodrošina būvniecības procesu atbilstību noteiktajiem standartiem un noteikumiem, palīdzot novērst defektus.
- **CPM (būvniecības projektu vadība)** - ietver būvniecības procesu plānošanu, koordinēšanu un kontroli.
- **CAFM (Computer-Aided Facility Management)** - ēku vadības un uzturēšanas sistēmas.
- Lai optimizētu materiālu un informācijas plūsmu starp piegādātājiem un būvlaukumu, ir nepieciešama SCM (piegādes ķedes pārvaldība).
- **EPM (Enterprise Performance Management)** - mērķis ir uzlabot biznesa procesus un veikspēju.
- **AMS (aktīvu pārvaldības programmatūra)** - izmanto, lai optimizētu iekārtu un infrastruktūras izmantošanu, pārvaldību un uzturēšanu visā aktīvu dzīves ciklā.
- **RPM (Nekustamā īpašuma pārvaldība)** - ietver uzdevumus un procesus, kas saistīti ar ēku un zemes, kā arī saistīto resursu un aktīvu pārvaldību un ekspluatāciju.



Attēls 1.2-4 Sistēmu savstarpējā savienojamība, kas savieno uzņēmuma procesus ar informācijas plūsmu starp dažādiem departamentiem.

- **CAE (Computer-Aided Engineering)** - datorizētā inženierija, kas ietver skaitlošanas un simulācijas sistēmas, piemēram, galīgo elementu analīzi (FEA) un skaitlošanas šķidrumu dinamiku (CFD).
- **CFD (Computational Fluid Dynamics)** - aprēķinu šķidrumu dinamika, šķidrumu un gāzu plūsmu modelēšana. CAE apakškategorija.
- **CAPP (Datorizēta procesu plānošana)** - Datorizēta procesu plānošana. To izmanto maršrutu un procesu karšu izveidei.
- **CAM (Computer-Aided Manufacturing)** - datorizētā ražošana, CNC mašīnu vadības programmu ģenerēšana.
- **PDM (Product Data Management)** - produkta datu pārvaldība, tehniskās dokumentācijas glabāšanas un pārvaldības sistēma.
- **MES (Manufacturing Execution System)** ir reāllaika ražošanas procesu vadības sistēma.
- **PLM (Product Lifecycle Management)** - projekta elementa dzīves cikla pārvaldība, integrē PDM, CAPP, CAM un citas sistēmas, lai nodrošinātu pilnīgu produkta kontroli no izstrādes līdz likvidācijai.

Šīs un daudzas citas sistēmas, kas ietver dažādus programmatūras risinājumus, ir kļuvušas par mūsdienu būvniecības nozares neatņemamu sastāvdaļu ( 1.2-4. attēls). Šādu sistēmu būtība ir specializētas datubāzes ar intuitīvām saskarnēm, kas nodrošina efektīvu informācijas ievadīšanu, apstrādi un analīzi visos projektēšanas un būvniecības posmos. Digitālo rīku savstarpēja integrācija ne tikai palīdz optimizēt darba procesus, bet arī būtiski uzlabo lēmumu precīzitāti, kas pozitīvi ietekmē projekta īstenošanas termiņus un kvalitāti.

Taču pusē gadījumu integrācija nenotiek. Saskaņā ar statistikas datiem tikai katra otrā lietojumprogramma vai sistēma ir integrēta ar citiem risinājumiem [17]. Tas norāda uz digitālās vides pastāvīgu sadrumstalotību un uzsver nepieciešamību izstrādāt atvērtus standartus un vienotas saskarnes, lai nodrošinātu visaptverošu informācijas apmaiņu būvniecības projekta ietvaros.

Viens no galvenajiem integrācijas izaicinājumiem mūsdienu uzņēmumos joprojām ir digitālās sistēmas lielā sarežģītība un prasības lietotāju kompetencei, kas nepieciešama efektīvai informācijas meklēšanai un interpretācijai. Katras uzņēmumā ieviestās sistēmas atbalstam tiek veidota speciālistu komanda, kuru vada galvenais vadītājs ( 1.2-2. attēls).

Galvenajam sistēmas pārvaldniekam ir izšķiroša loma datu plūsmas pareizā virzībā un viņš ir atbildīgs par galīgās informācijas kvalitāti, tāpat kā pirms tūkstošiem gadu pirmie pārvaldnieki bija atbildīgi par cipariem, kas bija uzrakstīti uz papirusa vai māla plāksnēm.

Lai atšķirīgas informācijas plūsmas pārvērstu par pārvaldības rīku, būtiska ir spēja sistemātiski integrēt un pārvaldīt datus. Šajā arhitektūrā vadītājiem ir jādarbojas kā vienota tīkla elementiem - kā micēlijam, kas savieno atsevišķas uzņēmuma daļas vienotā dzīvā organismā, kurš spēj pielāgoties un attīstīties.

## **Uzņēmuma micēlijs: kā dati savienojas ar uzņēmējdarbības procesiem**

Datu integrēšanas process lietojumprogrammās un datubāzēs ir balstīts uz informācijas apkopošanu

no dažādiem avotiem, tostarp no dažādām nodaļām un speciālistiem (1.2-4. attēls). Speciālisti meklē attiecīgos datus, apstrādā tos un pārnes uz savām sistēmām un lietojumprogrammām tālākai izmantošanai.

Katra uzņēmuma sistēma, kas sastāv no rīku, tehnoloģiju un datubāzu kopuma, ir zināšanu koks, kas saknējas vēsturisko datu augsnē un aug, lai dotu jaunus augļus gatavu risinājumu veidā: dokumenti, aprēķini, tabulas, grafiki un informācijas paneli (1.2-5. attēls). Sistēmas uzņēmumā, līdzīgi kokiem noteiktā meža pudurī, mijiedarbojas un sazinās savā starpā, veidojot sarežģītu un labi strukturētu sistēmu, ko atbalsta un pārvalda pieredzējuši vadītāji.

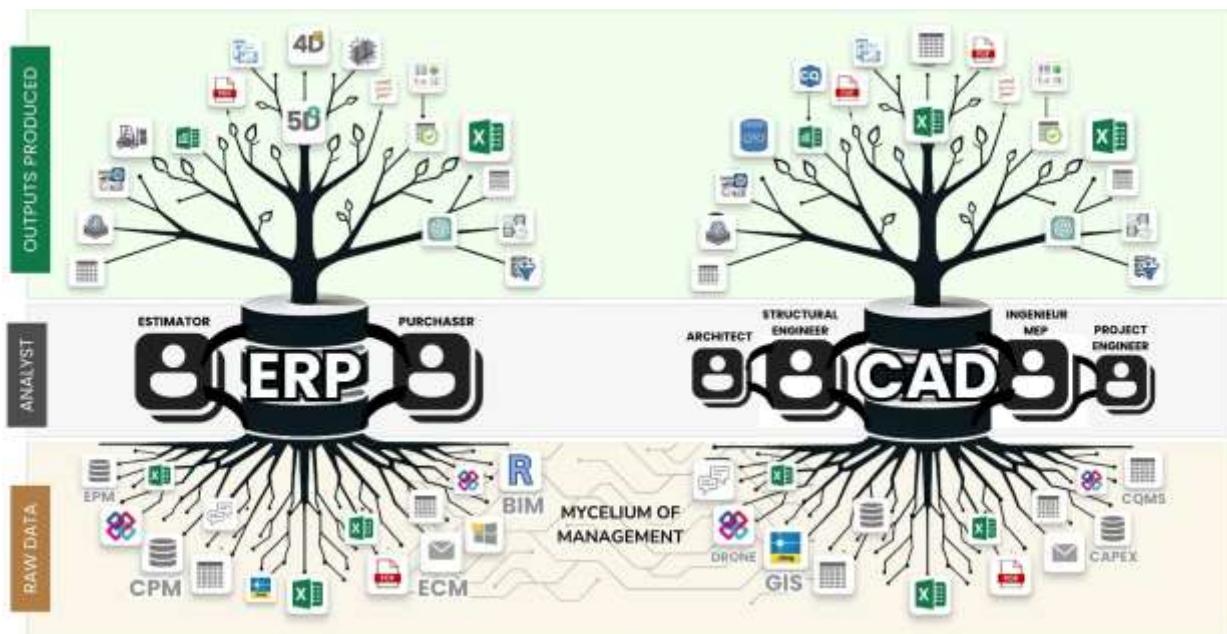
Uzņēmuma informācijas iegūšanas un pārsūtīšanas sistēma darbojas kā sarežģīts meža tīkls, kas sastāv no kokiem (sistēmām) un micēlijā sēnēm (vadītājiem), kuri darbojas kā vadītāji un pārstrādātāji, nodrošinot informācijas pārsūtīšanu un plūsmu uz pareizajām sistēmām. Tas palīdz uzturēt veselīgu un efektīvu datu plūsmu un izplatīšanu uzņēmumā.

Eksperti, līdzīgi kā saknes, uzsūc neapstrādātus datus projekta sākotnējā posmā, pārvēršot tos par barības vielām uzņēmuma ekosistēmai. Datu un satura pārvaldības sistēmas (1.2-4. attēls - ERP, CPM, BIM u. c.) darbojas kā spēcīgas informācijas maģistrāles, pa kurām šīs zināšanas cirkulē visos uzņēmuma līmeņos.

Līdzīgi kā dabā, kur katram ekosistēmas elementam ir sava loma, arī uzņēmuma uzņēmēdarbības vidē katrs procesa dalībnieks - no inženiera līdz analītiķim - veicina informācijas vides izaugsmi un auglību. Šie sistēmiskie "datu koki" (1.2-5. attēls) ir ne tikai informācijas apkopošanas mehānismi, bet arī konkurences priekšrocība, kas nodrošina uzņēmuma ilgtspēju.

Meža ekosistēmas pārsteidzoši precīzi atspoguļo to, kā ir organizētas digitālās korporatīvās struktūras. Līdzīgi kā meža daudzpakāpu struktūra - no meža zemsedzes līdz koku galotnēm - korporatīvā pārvaldība piešķir uzdevumus atbildības līmeņiem un funkcionālajiem departamentiem.

Dziļas un sazarotas koku saknes nodrošina izturību un piekļuvi barības vielām. Līdzīgi arī stabila organizatoriskā struktūra un stabili procesi darbam ar kvalitatīviem datiem atbalsta visu uzņēmuma informācijas ekosistēmu, veicinot tās ilgtspējīgu izaugsmi un attīstību pat tirgus nestabilitātes un krīžu periodos.



**Attēls 1.2-5 Datu integrēšana, izmantojot dažadas sistēmas, ir kā micēlijs, kas savieno vadītajus un speciālistus vienotā informācijas tīklā.**

Mūsdieni izpratne par mērogu uzņēmējdarbībā ir attīstījusies. Mūsdienās uzņēmuma vērtību nosaka ne tikai tā redzamā daļa - "kroni" gala dokumentu un pārskatu veidā -, bet arī kvalitatīvi apkopoto un sistemātiski apstrādāto datu "sakņu sistēmas" dzīlums. Jo vairāk informācijas var savākt un apstrādāt, jo lielāka kļūst uzņēmuma vērtība. Uzņēmumi, kas metodiski uzkrāj jau apstrādātu datu "kompostu" un spēj no tā iegūt noderīgas atziņas, iegūst stratēģiskas priekšrocības

Vēsturiskā informācija kļūst par jauna veida kapitālu, kas nodrošina izaugsmi, procesu optimizāciju un konkurences priekšrocības. Uz datiem balstītā pasaulei uzvar nevis tas, kam ir vairāk, bet gan tas, kurš zina vairāk.

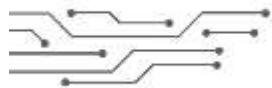
Būvniecības nozarē tas nozīmē pāreju uz reāllaika projektu pārvaldību, kur visi procesi - no projektēšanas un iepirkuma līdz darbuzņēmēju koordinēšanai - balstīsies uz attiecīgiem, katru dienu atjauninātiem datiem. Informācijas integrācija no dažādiem avotiem (ERP -sistēmas, CAD -modeļi, sensori IoT būvlaukumos, RFID) ļaus izstrādāt precīzākas prognozes, ātri reāgēt uz izmaiņām un izvairīties no kavēšanās, ko izraisa aktuālu datu trūkums.

Saskaņā ar McKinsey & Company "Data-Driven Enterprise 2025" (McKinsey & Company®, 2022 [18]) veiksmīgie nākotnes uzņēmumi paļausies uz datiem visos galvenajos savas darbības aspektos, sākot no stratēģiskiem lēmumiem līdz operatīvai mijiedarbībai.

Dati vairs nebūs tikai analīzes rīks un kļūs par visu uzņēmējdarbības procesu neatņemamu sastāvdaļu, nodrošinot pārredzamību, kontroli un pārvaldības automatizāciju. Uz datiem balstīts gājiens ļaus organizācijām līdz minimumam samazināt cilvēciskā faktora ietekmi, samazināt darbības riskus un palielināt lēmumu pieņemšanas pārredzamību un efektivitāti.

21. gadsimts apgriež ekonomikas paradigmu otrādi: ja agrāk naftu dēvēja par "melno zeltu", jo tā spēja

darbināt mašīnas un transportu, tad šodien, saspiesti laika spiediena apstākļos, vēsturiskie dati kļūst par jaunu stratēģisko resursu, kas darbina nevis mašīnas, bet gan lēmumu pieņemšanas algoritmus, kuri vadīs uzņēmējdarbību.



## NODAĻA 1.3.

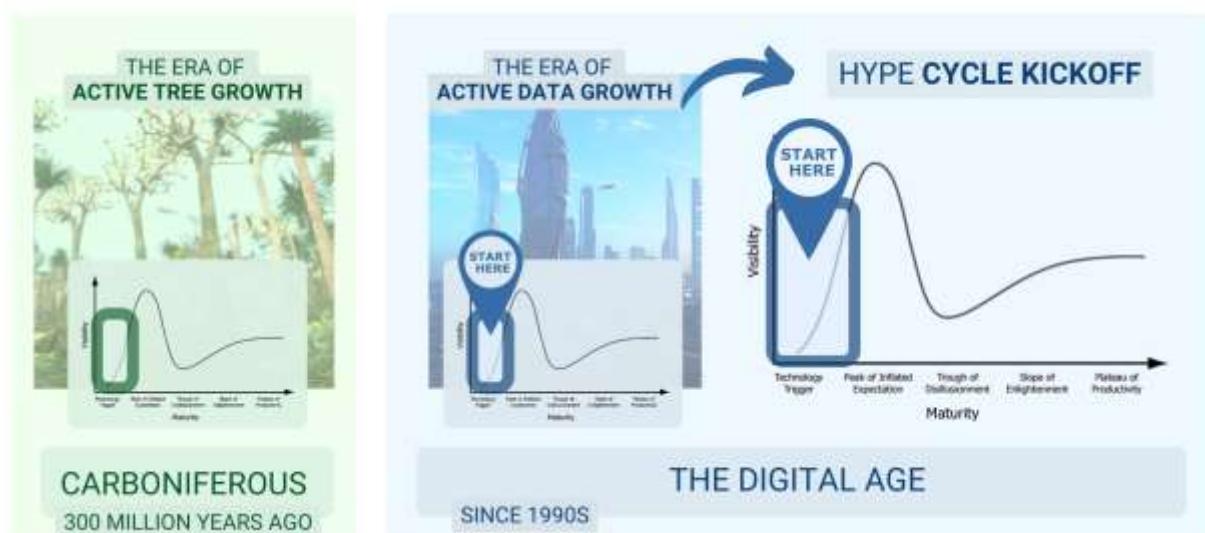
### DIGITĀLĀ REVOLŪCIJA UN DATU EKSPLOZIJA

#### Datu buma sākums kā evolūcijas vilnis

Būvniecības nozarē notiek vēl nebijis informācijas sprādziens. Ja uzņēmējdarbību iedomājamies kā zināšanu koku (1.2-5. attēls), ko papildina dati, tad pašreizējo digitalizācijas posmu var salīdzināt ar straujo veģetācijas attīstību karbona periodā, kad Zemes biosfēru pārveidoja strauja biomasa uzkrāšanās (1.3-1. attēls).

Līdz ar globālo digitalizāciju informācijas apjoms būvniecības nozarē katru gadu dubultojas. Modernās tehnoloģijas jauj datus vākt fonā, analizēt reāllaikā un izmantot tādā mērogā, kas vēl nesen šķita neiespējami.

Saskaņā ar Mūra likumu, ko formulēja Gordons Mūrs (Intel® līdzdibinātājs), integrālo shēmu blīvums un sarežģītība, kā arī apstrādājamo un uzglabājamo datu apjoms dubultojas aptuveni ik pēc diviem gadiem [19].



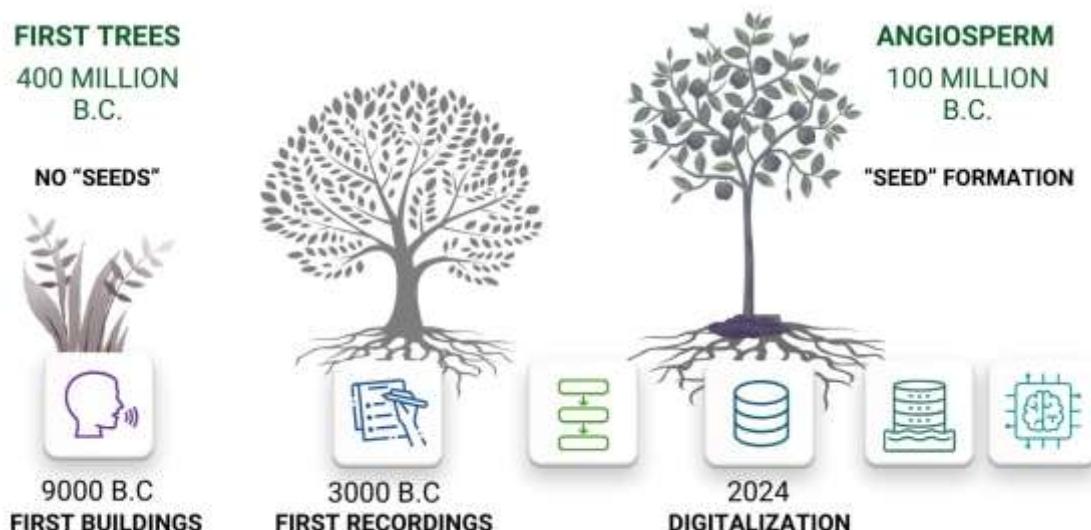
Attēls 1.3-1 Digitalizācijas sākums ir izraisījis eksponenciālu datu apjoma pieaugumu, līdzīgi kā veģetācijas uzplaukums ogļu laikmetā.

Lai gan senās megalītiskās celtnes, piemēram, Göbekli Tepe (Turcija), nav atstājušas dokumentētas zināšanas, kas būtu piemērotas atkārtotai izmantošanai, mūsdienās digitālās tehnoloģijas dod iespēju uzkrāt un atkārtoti izmantot informāciju. To var salīdzināt ar evolūcijas pāreju no sporu augiem uz sēklu augiem (angiospermiem): sēklu parādišanās veicināja dzīvības plašu izplatību uz planētas. (1.3-2. attēls).

Līdzīgi, dati no iepriekšējiem projektiem kļūst par sava veida "digitālajām sēklām" - DNS zināšanu nesējiem, ko var paplašināt un izmantot jaunos projektos un produces. Mūsdienu mākslīgā intelekta rīku - mašīnmācīšanās un lielo valodas modeļu (LLM), piemēram, ChatGPT, LlaMa, Mistral, Claude,

DeepSeek, QWEN, Grok - parādīšanās ļauj datus automātiski iegūt, interpretēt un izmantot jaunos kontekstos

Līdzīgi kā sēklas revolucionāri veicināja dzīvības izplatīšanos uz sākotnēji nedzīvas planētas, "datu sēklas" kļūst par pamatu jaunu informācijas struktūru un zināšanu automātiskai rašanās procesam, ļaujot digitālajām ekosistēmām patstāvīgi attīstīties un pielāgoties mainīgajām lietotāju prasībām.



Attēls 1.3-2 Digitālajām "datu sēklām" ir tāda pati evolūcijas loma kā angiospermiem - ziedošiem augiem, kas pārveidoja Zemes ekosistēmu.

Mēs atrodamies uz jaunas ēras sliekšņa būvniecībā, kur datu eksplozija un aktīva "datu sēku" - strukturētas informācijas no iepriekšējiem un pašreizējiem projektiem - izplatīšana veido nozares digitālās nākotnes pamatu. To "apputeksnēšana", izmantojot lielo datu valodas modeļus (LLM), ļauj mums ne tikai vērot digitālās pārmaiņas, bet arī aktīvi piedalīties pašmācības un adaptīvu ekosistēmu veidošanā. Tā nav evolūcija - tā ir digitālā revolūcija, kurā dati kļūst par jaunās realitātes galveno pamatelementu

*Būvniecības nozarē krasī pieaug datu apjoms, jo visā būvprojekta dzīves ciklā tiek iegūta informācija no dažādām nozarēm. Šī milzīgā datu uzkrāšanās ir virzījusi būvniecības nozari uz lielo datu laikmetu [20].*

- Prof. Hang Yang, Būvniecības inženierzinātņu un arhitektūras fakultāte, Wuhan Tehnoloģiju universitāte, Wuhan.,

Datu apjoma pieaugums informācijas laikmetā atgādina evolūcijas procesus dabā: tāpat kā mežu attīstība mainīja senās planētas ainavu, pašreizējais informācijas sprādziens maina visas būvniecības nozares ainavu.

## Mūsdienu uzņēmumā ģenerēto datu apjoms

Pēdējo divu gadu laikā ir radīti 90% no visiem pasaule esošajiem datiem [21]. Sākot ar 2023. gadu, katrs cilvēks, ieskaitot būvniecības nozares profesionāļus, rada aptuveni 1,7 megabaitus datu sekundē [22], un kopējais datu apjoms pasaule 2023. gadā sasniegis 64 zetabaitus, un tiek prognozēts, ka 2025. gadā tas pārsniegs 180 zetabaitus jeb  $180 \times 10^{15}$  megabaitus [23].

Šim informācijas sprādzienam ir vēsturisks precedents - Johannesa Güttenberga drukas izgudrojums 15. gadsimtā. Tikai piecdesmit gadus pēc tās ieviešanas grāmatu skaits Eiropā dubultojās: dažu desmitgažu laikā tika iespiests tikpat daudz grāmatu, cik iepriekšējos 1200 gados bija radīts ar rokām [24]. Mūsdienās vērojams vēl straujāks pieaugums: datu apjoms pasaule dubultojas ik pēc trim gadiem.

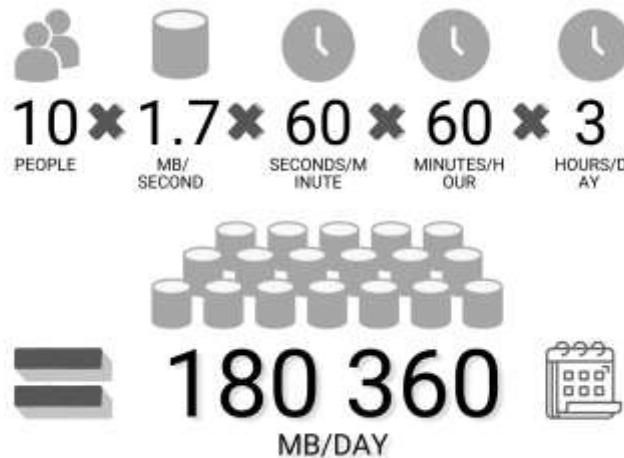
Nemot vērā pašreizējos datu pieauguma tempus, būvniecības nozarei tuvākajās desmitgadēs ir potenciāls radīt tik daudz informācijas, cik tā ir uzkrājusi visā savā līdzšinējā vēsturē



Attēls 1.3-3 Ikdienas datu glabāšana uzņēmuma serveros, ko veic katrs darbinieks, veicina pastāvīgu datu pieaugumu.

Mūsdienu būvniecības biznesa pasaule pat mazi uzņēmumi ik dienu ģenerē milzīgu daudzformāta informācijas apjomu, un pat neliela būvniecības uzņēmuma digitālā pēda var sasniegt desmitiem gigabaitu dienā - no modeļiem un rasējumiem līdz fotoattēliem un sensoriem uz vietas. Ja pieņemam, ka katrs tehnīķis vidēji rada aptuveni 1,7 MB datu sekundē, tas atbilst aptuveni 146 GB dienā jeb 53 TB gadā ( 1.3-3. attēls).

Ja 10 cilvēku komanda aktīvi strādā tikai 3 stundas dienā, dienā kumulatīvais radītās informācijas apjoms sasniedz 180 gigabaitus ( 1.3-4. attēls).



Attēls 1.3-4 Uzņēmums, kurā strādā 10 cilvēki, dienā ģenerē aptuveni 50-200 gigabaitus datu.

Pieņemot, ka 30% darba datu ir jauni (pārējie tiek pārrakstīti vai dzēsti), uzņēmums ar 10 darbiniekiem var radīt vairākus simtus gigabaitu jaunu datu mēnesī (faktiskie skaitli ir atkarīgi no uzņēmuma darbības veida)

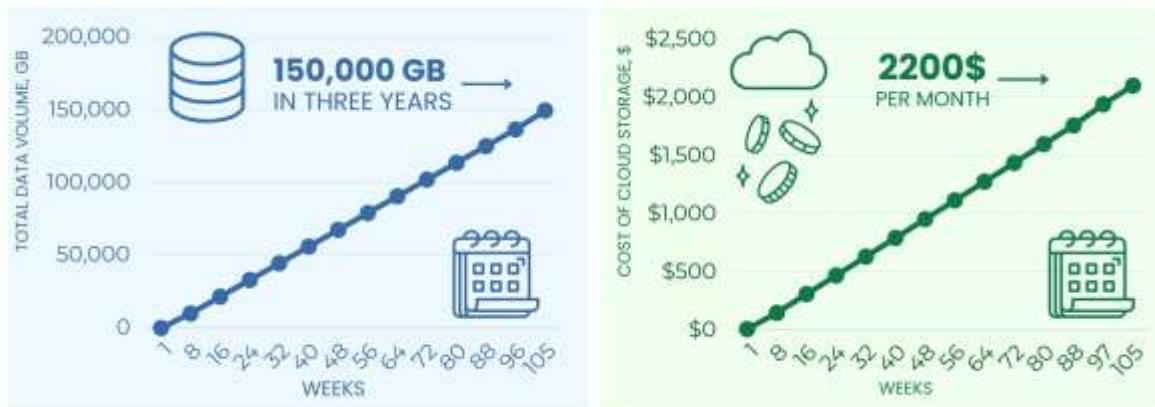
Tādējādi ir skaidrs, ka mēs ne tikai ģenerējam arvien vairāk datu, bet arī saskaramies ar pieaugošu vajadzību pēc to efektīvas pārvaldības, uzglabāšanas un ilgtermiņa pieejamības. Un, lai gan iepriekš dati varēja "gulēt" uz vietējiem serveriem bez maksas, digitālās transformācijas kontekstā arvien vairāk uzņēmumu sāk izmantot mākoņrisinājumus kā savas informācijas infrastruktūras pamatu.

### Datu glabāšanas izmaksas: ekonomiskais aspeks

Pēdējos gados arvien vairāk uzņēmumu izmanto mākoņpakuļojumus, lai ārpakalpojumos uzglabātu datus. Piemēram, ja uzņēmums pusi no saviem datiem izvieto mākonī, un vidējā cena ir 0,015 ASV dolāru par gigabaitu mēnesī, tā datu glabāšanas izmaksas katru mēnesi var palielināties par 10-50 ASV dolāriem [25].

Mazam uzņēmumam ar tipiskiem datu ģenerēšanas modeļiem mākoņdatu glabāšanas izmaksas var svārstīties no simtiem līdz potenciāli vairāk nekā tūkstoš dolāriem mēnesī (1.3-5. attēls) dažu gadu laikā, radot potenciāli ievērojamu finansiālu slogu.

Saskaņā ar Forrester pētījumu "Uzņēmumi, pieaugot sarežģītībai, nodod datu glabāšanu ārpakalpojumu sniedzējiem" [26], kurā aptaujāti 214 tehnoloģiju infrastruktūras lēmumu pieņēmēji. [26], kurā aptaujāti 214 tehnoloģiju infrastruktūras lēmumu pieņēmēji, vairāk nekā trešdaļa organizāciju izmanto ārpakalpojumus datu glabāšanai, lai tiktu galā ar pieaugošo datu apjomu un sarežģītību, un gandrīz divas trešdaļas uzņēmumu dod priekšroku abonēšanas modelim.



Attēls 1.3-5 Datu pārvietošana uz mākonī var palielināt ikmēneša glabāšanas izmaksas līdz pat 2000 ASV dolāru pat uzņēmumam, kurā ir tikai 10 darbinieki.

Situāciju vēl vairāk sarežģī paātrinātā mākoņtehnoloģiju, piemēram, CAD (BIM), CAFM, PMIS un ERP - sistēmu, ieviešana, kas vēl vairāk palielina datu uzglabāšanas un apstrādes izmaksas. Rezultātā uzņēmumi ir spiesti meklēt veidus, kā optimizēt izmaksas un samazināt atkarību no mākoņpakaļpojumu sniedzējiem.

Kopš 2023. gada, aktīvi attīstoties lielajiem valodas modeļiem (LLM), ir sākušas mainīties pieejas datu glabāšanai. Arvien vairāk uzņēmumu domā par savu datu kontroles atgūšanu, jo klūst drošāk un izdevīgāk apstrādāt informāciju savos serveros.

Šajā kontekstā priekšplānā izvirzās tendence atteikties no mākoņdatošanas un tikai nepieciešamo datu apstrādes par labu uzņēmumu LLM un AI risinājumu vietējai izvietošanai. Kā vienā no intervijām [27] norādīja Microsoft izpilddirektors, tā vietā, lai dažādu uzdevumu veikšanai izmantotu vairākas atsevišķas lietojumprogrammas vai mākoņtehnoloģiju SaaS risinājumus, AI aģenti pārvaldīs procesus datubāzēs, automatizējot dažādu sistēmu funkcijas.

[...] vecā pieeja šim [datu apstrādes] jautājumam bija šāda: ja atceramies, kā dažādas biznesa lietojumprogrammas veica integrāciju, tās izmantoja savienotājus. Uzņēmumi pārdeva šo savienotāju licences, un ap to tika izveidots uzņēmējdarbības modelis. SAP [ERP] ir viens no klasiskajiem piemēriem: jūs varējāt piekļūt SAP datiem tikai tad, ja jums bija pareizais savienotājs. Tāpēc man šķiet, ka kaut kas līdzīgs parādīsies arī attiecībā uz [AI] aģentu mijiedarbību [...]. Vismaz mūsu pieeja ir šāda: es domāju, ka [mākslīgā intelekta] aģentu laikmetā, iespējams, sabruks koncepcija par biznesa lietojumprogrammu pastāvēšanu. Jo, ja par to padomā, tie būtībā ir datu bāzes ar virkni biznesa loģikas

- Satya Nadella, Microsoft izpilddirektors, intervija ar kanālu BG2, 2024. gads. [28]

Šajā paradigmā uz datiem balstīta LLM pieeja pārsniedz klasiskās sistēmas. Mākslīgais intelekts klūst par starpnieku starp lietotāju un datiem ( 2.2-3. attēls, 2.2-4. attēls), novēršot nepieciešamību pēc

vairākām starposma saskarnēm un palielinot uzņēmējdarbības procesu efektivitāti. Plašāk par šo pieeju darbam ar datiem mēs runāsim nodaļā "Haosa pārvēršana kārtībā un sarežģītības samazināšana".

Lai gan nākotnes arhitektūra vēl tikai veidojas, uzņēmumi jau saskaras ar iepriekš pieņemto lēmumu sekām. Pēdējo desmitgažu masveida digitalizācija, ko pavadīja atšķirīgu sistēmu ieviešana un nekontrolēta datu uzkrāšana, ir radījusi jaunu problēmu - informācijas pārslodzi.

## Datu uzkrāšanas robežas: no masas līdz nozīmei

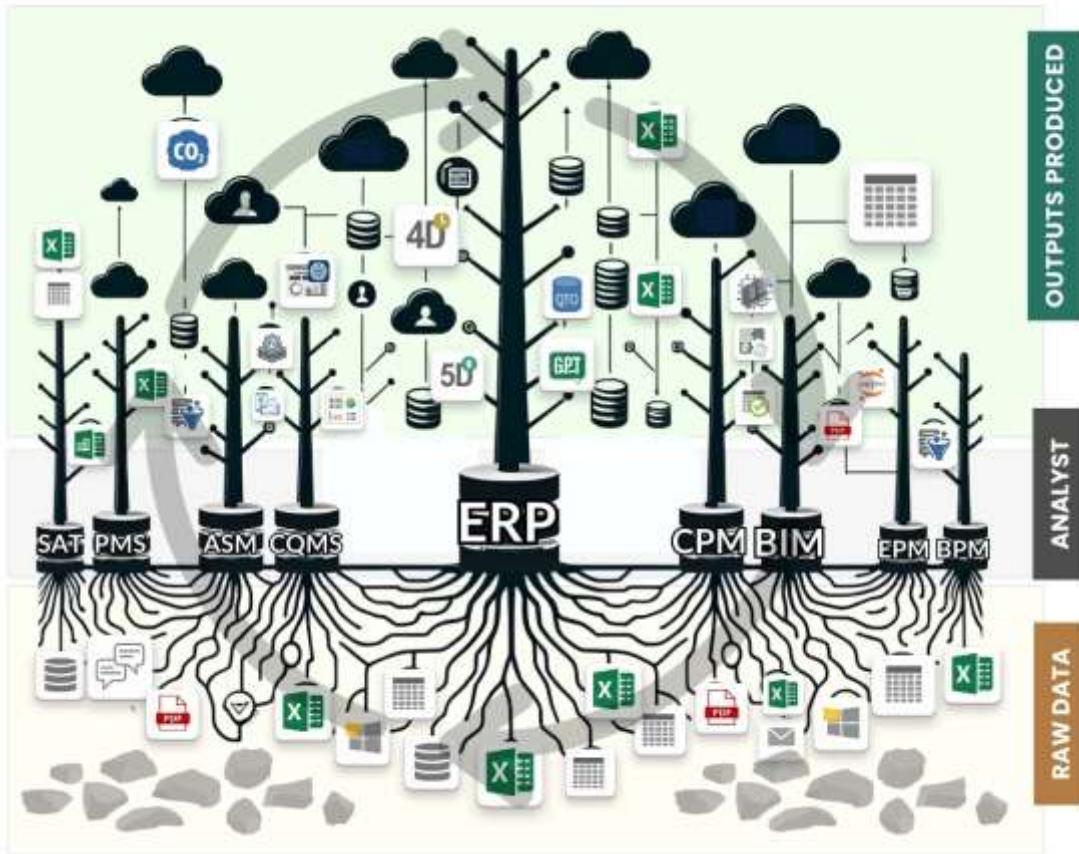
Mūsdienīgas uzņēmumu sistēmas veiksmīgi attīstās un darbojas kontrolētas izaugsmes apstākļos, kad datu apjoms un lietojumprogrammu skaits ir līdzsvarā ar IT nodaļu un vadītāju iespējām. Tomēr pēdējās desmitgadēs digitalizācija ir izraisījusi nekontrolējamu datu apjoma un sarežģītības pieaugumu, kas uzņēmumu informācijas ekosistēmā ir radījis pārsātinājuma efektu.

Mūsdienās serveri un krātuves ir pakļauti vēl nepieredzētam neapstrādātas un daudzformāta informācijas pieplūdumam, kas nav paguvusi pārvērsties kompostā un ātri vien klūst nebūtiska. Ierobežotie uzņēmuma resursi nespēj tikt galā ar šo pārplūdumu, un dati uzkrājas izolētos silosos (tā dēvētajos "silosos"), no kuriem noderīgas informācijas iegūšanai nepieciešama manuāla apstrāde.

Tā rezultātā, līdzīgi kā mežs, kas aizaudzis ar efeju un apaugis ar pelējumu, mūsdienu uzņēmumu vadības sistēmas bieži cieš no informācijas pārslodzes. Tā vietā, lai uzņēmuma ekosistēmas centrā uzturētu informācijas humusu, veidojas izolētas dažāda formāta datu zonas, kas neizbēgami novēd pie uzņēmējdarbības procesu kopējās efektivitātes samazināšanās.

Pēdējo 40 gadu ilgajam eksponenciālā datu apjoma pieauguma periodam neizbēgami sekos piesātināšanās un tai sekojošais atdzišanas posms. Kad datu glabāšana sasniegts savas robežas, notiks kvalitatīva pārmaiņa: dati vairs nebūs tikai glabāšanas objekts, bet gan stratēģisks resurss.

Attīstoties mākslīgajam intelektam un mašīnmācībai, uzņēmumiem ir iespēja samazināt informācijas apstrādes izmaksas un pāriet no kvantitatīvas izaugsmes uz kvalitatīvu datu izmantošanu. Nākamajā desmitgadē būvniecības nozarei būs jānovirza uzmanība no arvien lielāka datu apjoma radīšanas uz to struktūras, integritātes un analītiskās vērtības nodrošināšanu.



Attēls 1.3-6 Izolēti datu avoti kavē informācijas apmaiņu starp datu sistēmām.

Galvenā vērtība vairs nav informācijas apjomā, bet gan spējā to automātiski interpretēt un pārvērst lietišķās zināšanās, kas noder vadības lēmumu pieņemšanai. Lai dati kļūtu patiesi noderīgi, tie ir pareizi jāpārvalda: jāapkopo, jāpārbauda, jāstrukturē, jāuzglabā un jāanalizē konkrētu uzņēmējdarbības uzdevumu kontekstā.

Datu analīzes process uzņēmumā ir līdzīgs koku dzīves un pūšanas ciklam mežā un jaunu, spēcīgu un jaunu koku augšanai: nobrieduši koki iet bojā, sadalās un kļūst par augsnī jaunam augumam. Pabeigtie un pabeigtie procesi, kad tie ir pabeigti, kļūst par daļu no uzņēmuma informācijas ekosistēmas, galu galā kļūstot par informācijas humusu, kas veicina jaunu sistēmu un datu turpmāku izaugsmi.

Tomēr praksē šis cikls bieži vien tiek pārtraukts. Tā vietā, lai organiski atjaunotos, veidojas daudzslāņains haoss, līdzīgi kā ģeoloģiskajos slāņos, kur jaunas sistēmas tiek uzklātas uz vecajām bez dziļākas integrācijas un strukturēšanas. Rezultātā rodas atšķirīgi informācijas "silosi", kas kavē zināšanu apriti un sarežģī datu pārvaldību.

## Nākamie soļi: no datu teorijas līdz praktiskām pārmaiņām

Datu attīstība būvniecībā ir ceļojums no māla plāksnēm līdz mūsdienu moduļu platformām. Šodienas

uzdevums ir nevis apkopot informāciju, bet gan izveidot sistēmu, kas atšķirīgus un daudzveidīgus datus pārvērš stratēģiskā resursā. Neatkarīgi no tā, vai esat uzņēmuma vadītājs vai inženieris, izpratne par datu vērtību un to, kā ar tiem strādāt, būs viena no galvenajām prasmēm nākotnē.

Apkopojojot šo daļu, ir vērts uzsvērt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās pieejas ikdienas uzdevumos:

#### ■ Informācijas plūsmu personiskā audita veikšana

- Sastādiet sarakstu ar visām sistēmām un lietojumprogrammām, ar kurām ikdienā strādājat.
- Atzīmējiet, kur pavadāt visvairāk laika, meklējot vai vēlreiz pārbaudot datus.
- Nosakiet galvenos informācijas avotus
- Analizējiet savu pašreizējo lietojumprogrammu vidi, lai konstatētu, vai nav dublējošās un dublētās funkcijas.

#### ■ Centieties pāri procesos pa analītiskā brieduma līmeniem.

- Uzdevumus sāciet ar aprakstošo analīzi (kas notika?)
- Pakāpeniski ieviest diagnostiku (kāpēc tas notika?)
- Padomājiet par to, kā procesos var pāriet uz prognozējošo (kas notiks?) un preskriptīvo (ko darīt?) analītiku.

#### ■ Darba datu strukturēšanas sākšana

- Ievietojiet vienotu sistēmu bieži darbā izmantoto failu un mapju nosaukumu piešķiršanai.
- Izveidojiet veidnes bieži izmantotiem dokumentiem un pārskatiem.
- Regulāri arhivējiet pabeigtos projektus ar skaidru struktūru.

Pat ja nevarat mainīt visu savas komandas vai uzņēmuma informācijas infrastruktūru, sāciet ar saviem procesiem un nelieliem uzlabojumiem ikdienas darbā. Atcerieties, ka patiesā datu vērtība nav to apjomā, bet gan spējā no tiem iegūt noderīgas atziņas. Pat nelielām, bet pareizi strukturētām un analizētām datu kopām var būt būtiska ietekme, ja tās integrē lēmumu pieņemšanas procesos.

Turpmākajās grāmatas daļās mēs pāriesim pie konkrētām metodēm un rīkiem darbam ar datiem, aplūkosim veidus, kā nestukturētu informāciju pārveidot strukturētās kopās, izpētīsim analītikas automatizācijas tehnoloģijas un detalizēti aprakstīsim, kā izveidot efektīvu analītikas ekosistēmu būvniecības uzņēmumā.





## II DAĻA

# KĀ BŪVNIECĪBAS NOZARE SLĪGST DATU HAOSĀ

Otrā daļa ir veltīta kritiskai analīzei par problēmām, ar kurām saskaras būvniecības uzņēmumi, strādājot ar pieaugošiem datu apjomiem. Šīki aplūkotas informācijas sadrumstalotības sekas un "datu "silosu" fenomens, kas kavē efektīvu lēmumu pieņemšanu. Tieka pētītas HiPPO -pieejas (Highest Paid Person's Opinion) problēmas un tās ietekme uz vadības lēmumu kvalitāti būvniecības projektos. Novērtēta dinamisko uzņēmējdarbības procesu un to pieaugošās sarežģītības ietekme uz informācijas plūsmu un darbības efektivitāti. Tieka sniegti konkrēti piemēri, kā pārmērīga sistēmu sarežģītība palielina izmaksas un samazina organizāciju elastību. Īpaša uzmanība tiek pievērsta ierobežojumiem, ko rada patentēti formāti, un atvērto standartu izmantošanas perspektīvām būvniecības nozarē. Tieka izklāstīta koncepcija par pāreju uz mākslīgā intelekta un LLM balstītām programmatūras ekosistēmām, kas samazina pārmērīgu sarežģību un tehniskos šķēršļus.

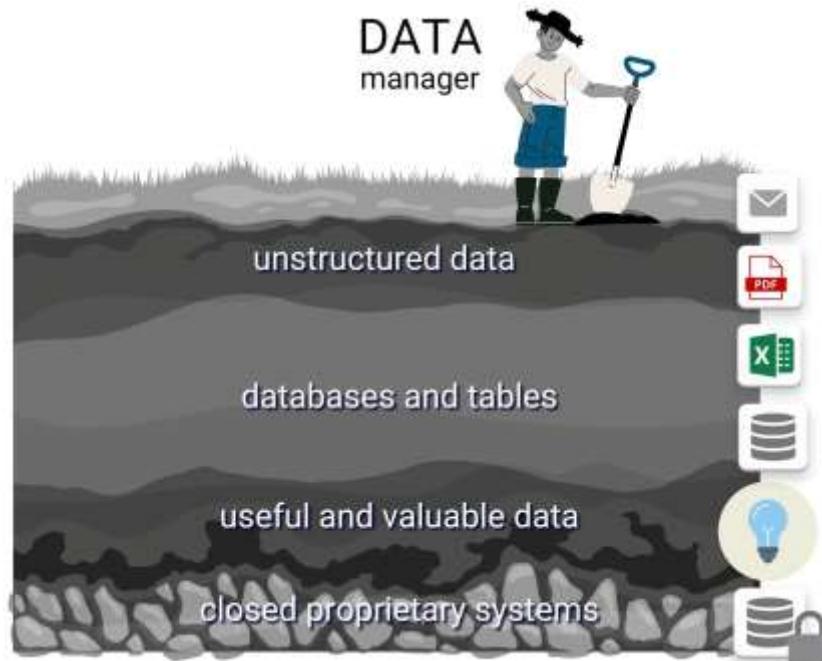
## NODĀLA 2.1.

### DATU SADRUMSTALOTĪBA UN "SILOS".

#### **Jo vairāk rīku, jo efektīvāks bizness?**

No pirmā acu uzmetiena var šķist, ka vairāk digitālo rīku nodrošina lielāku efektivitāti. Tomēr praksē tas tā nav. Ar katru jaunu risinājumu, neatkarīgi no tā, vai tas ir mākoņpakaļpojums, mantota sistēma vai kārtējais Excel pārskats, uzņēmums pievieno vēl vienu slāni savā digitālajā ainavā - slāni, kas bieži vien nav integrēts ar pārējiem ( 2.1-1. attēls).

Datus var salīdzināt ar oglēm vai naftu: tiem ir vajadzīgi gadi, lai uzkrātos, tos sablīvē haosa, kļūdu, nestukturētu procesu un aizmirstu formātu slāni. Lai no tiem iegūtu patiesi noderīgu informāciju, uzņēmumiem burtiski ir jāizlaužas cauri novecojušu risinājumu un digitālā trokšņa slāniem.



**Attēls 2.1-1 Daudzveidīgi dati veido sadalītus slāņus - pat "zelta" atziņas tiek pazaudētas ģeoloģiskajās sistēmiskās sarežģītības klintīs.**

Katra jauna lietojumprogramma aiz sevis atstāj pēdas: failu, tabulu vai veselu izolētu "silosu" serverī. Viens slānis ir māls (novecojuši un aizmirsti dati), otrs - smiltis (atšķirīgas tabulas un atskaites), bet trešais - granīts (slēgti patentēti formāti, kurus nevar integrēt). Laika gaitā uzņēmuma digitālā vide arvien vairāk līdzinās nekontrolētas informācijas uzkrāšanās rezervuāram, kur vērtība tiek zaudēta dziļi uzņēmuma serveros.

Ar katru jaunu projektu un katru jaunu sistēmu sarežģītāka kļūst ne tikai infrastruktūra, bet arī ceļš uz noderīgiem un kvalitatīviem datiem. Lai nonāktu līdz vērtīgajam "akmenim", ir nepieciešama dziļa

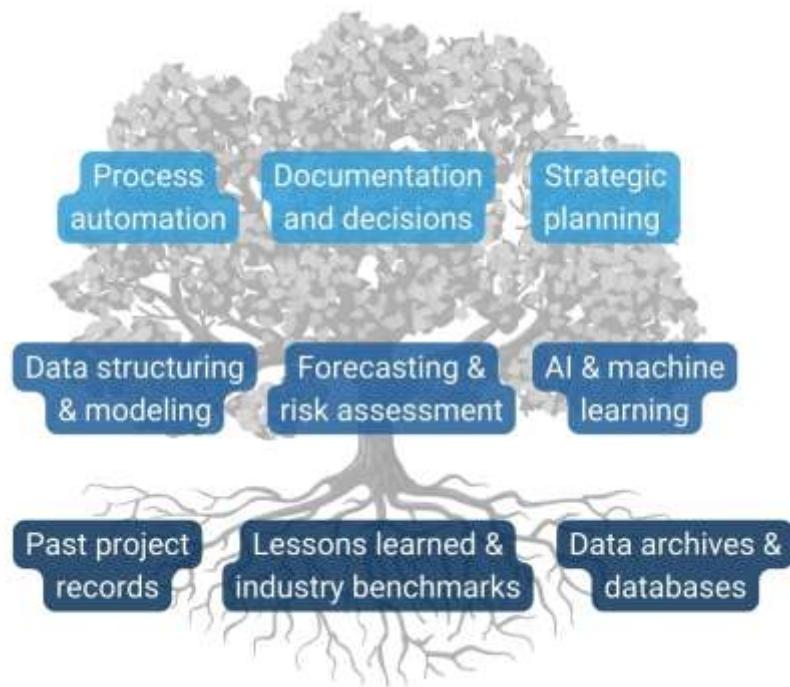
attīrišana, informācijas strukturēšana, "šķeldošana", grupēšana jēgpilnos gabalos un stratēģiski svarīgu ieskatu iegūšana, izmantojot analītiku un datu modelēšanu.

*Dati ir vērtīga lieta, un tie kalpos ilgāk nekā pašas sistēmas [kas apstrādā datus] [29].*

- Tims Berners-Lī, globālā tīmekļa tēvs un pirmās tīmekļa vietnes autors.

Pirms dati kļūst par "vērtīgu lietu" un uzticamu pamatu lēmumu pieņemšanai, tie ir rūpīgi jāsagatavo. Tieši pareiza pirmapstrāde ir tā, kas pārvērš atšķirīgus datus strukturētā pieredzē, noderīgā informācijas humusā, kas pēc tam kļūst par prognozēšanas un optimizācijas rīku.

Pastāv maldīgs priekšstats, ka, lai sāktu analizēt, ir nepieciešami pilnīgi tīri dati, taču praksē spēja strādāt ar netīriem datiem ir būtiska procesa daļa.



**Attēls 2.1-2 Dati ir uzņēmējdarbības sakņu sistēma un pamats, kas savukārt balstās uz lēmumu pieņemšanas procesiem.**

Tehnoloģijām turpinot attīstīties, arī jūsu uzņēmumam ir jāvirzās uz priekšu un jāmācās, kā no datiem radīt vērtību. Tāpat kā naftas un ogļu uzņēmumi veido infrastruktūru derīgo izrakteņu ieguvei, arī uzņēmumiem ir jāmācās, kā pārvaldīt jaunas informācijas plūsmu savos serveros un iegūt vērtīgas atzinjas no neizmantotiem, neformatētiem un novecojušiem datiem, pārvēršot tos par stratēģisku resursu.

Pirmais solis ir lauku izveide (datu noliktavas). Pat visjaudīgākie rīki neatrisina datu izolācijas un daudzformāta datu problēmu, ja uzņēmumi turpina darboties izolētās sistēmās. Ja dati eksistē atsevišķi viens no otra, savstarpēji nesaskaroties un nedaloties ar informāciju, uzņēmumi saskaras ar "datu silo" efektu. Tā vietā, lai izveidotu vienotu, konsekventu infrastruktūru, uzņēmumi ir spiesti tērēt resursus datu apvienošanai un sinhronizēšanai.

## Datu atslēgas un to ietekme uz uzņēmuma darbību

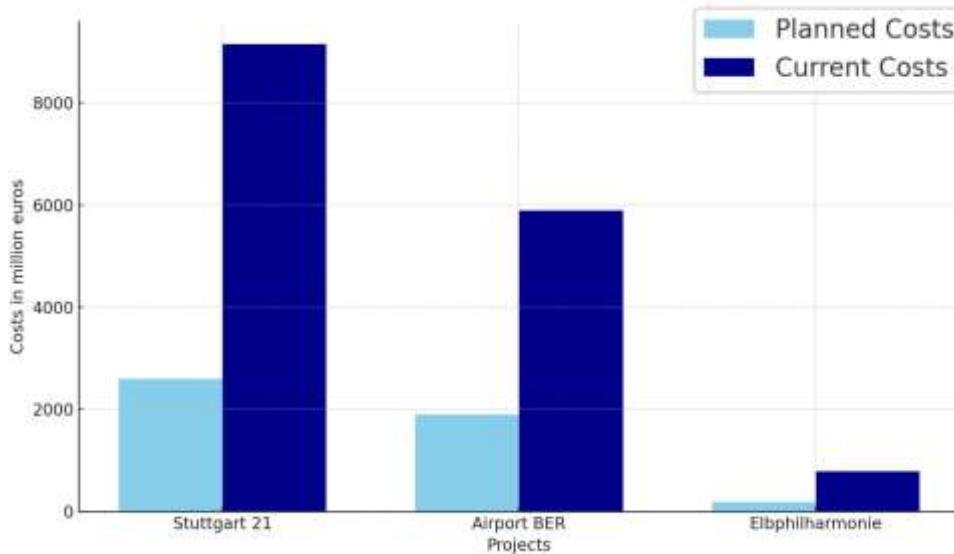
Iedomājieties, ka jūs būvējat dzīvojamu rajonu, bet katrai komandai ir sava projekts. Vieni būvē sienas, citi liek komunikācijas, bet vēl citi - ceļus, nepārbaudot cits citu. Rezultātā caurules neatbilst atverēm sienās, liftu šahtas neatbilst stāviem, un ceļi ir jānojauc un jāierīko no jauna.

Šī situācija nav tikai hipotētisks scenārijs, bet gan daudzu mūsdienu būvniecības projektu realitāte. Tā kā liels skaits ģenerāluzņēmēju un apakšuzņēmēju strādā ar dažādām sistēmām un nav vienota koordinācijas centra, process pārvēršas bezgalīgu apstiprinājumu, pārstrādāšanas un konfliktu virknē. Tas viss rada ievērojamu kavēšanos un vairākkārtējas projekta izmaksas.

Klasiska situācija būvlaukumā ir vienkārša: veidni ir gatavi, bet armatūras piegāde nav pienākusi laikā. Pārbaudot informāciju dažādās sistēmās, komunikācija ir aptuveni šāda:

- ⌚ Būvdarbu **vadītājs** 20. būvlaukumā raksta projekta vadītājam: "*Mēs esam pabeiguši veidot veidņus, bet kur ir armatūra?*"
- ⌚ **Projekta vadītājs** (PMIS) lepirkumu nodaļai: - "*Veidņi ir gatavi. Manā sistēmā [PMIS] rakstīts, ka armatūras stiegtrojums bija jāieved 18. aprīlī. Kur ir armatūra?*"
- ⌚ **Piegādes ķēdes speciālists** (ERP): - "*Mūsu ERP norāda, ka piegāde notiks 25. datumā*".
- ⌚ **Datu inženieris** vai IT departaments (atbildīgs par integrāciju): - PMIS datums ir 18., ERP - 25. datums. Starp ERP un PMIS nav saiknes ar pasūtījuma ID, tāpēc dati nav sinhronizēti. Šis ir tipisks informācijas trūkuma piemērs.
- ⌚ **Projekta vadītājs** ģenerāldirektoram - "*Armatūras piegāde kavējas, objekts stāv, un nav skaidrs, kas ir atbildīgs*".

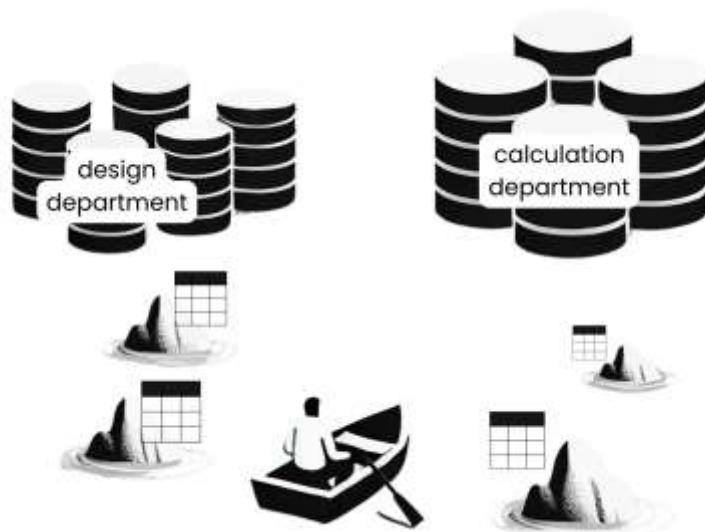
Incidenta cēlonis bija datu izolācija atšķirīgās sistēmās. Integrējot un apvienojot datu avotus, izveidojot vienotu informācijas repozitoriju un automatizējot ar ETL rīkiem (Apache NiFi, Airflow vai n8n), var likvidēt "silosus" starp sistēmām. Šīs un citas metodes un rīki tiks detalizēti aplūkoti turpmākajās grāmatas nodaļās.



Attēlā: 2.1-3 Vācijas lielāko infrastruktūras projektu plānoto un faktisko izmaksu salīdzinājums.

Tas pats notiek arī ar uzņēmumu sistēmām: vispirms tiek radīti atsevišķi risinājumi, un pēc tam to integrācijai un saskaņošanai ir jāiegulda milzīgi līdzekļi. Ja datu un saziņas modeļi būtu pārdomāti jau no paša sākuma, integrācija vispār nebūtu vajadzīga. Atsevišķi dati digitālajā pasaulē rada haosu, līdzīgi kā nesaskaņots būvniecības process.

Saskaņā ar KPMG 2023. gada pētījuma "Cue construction 4.0: Time to make or break" datiem tikai 36% uzņēmumu efektīvi koplieto datus starp departamenti, bet 61% saskaras ar nopietnām problēmām, ko rada izolēti datu "silosi" [30].



Attēlā 2.1-4 Gadu gadiem grūti savācam i datu uzkrājas izolētās datu glabātuvēs, riskējot, ka tie nekad netiks izmantoti.

Uzņēmuma dati tiek glabāti izolētās sistēmās kā atsevišķi koki, kas izkaisīti pa ainavu. Katrā no tiem

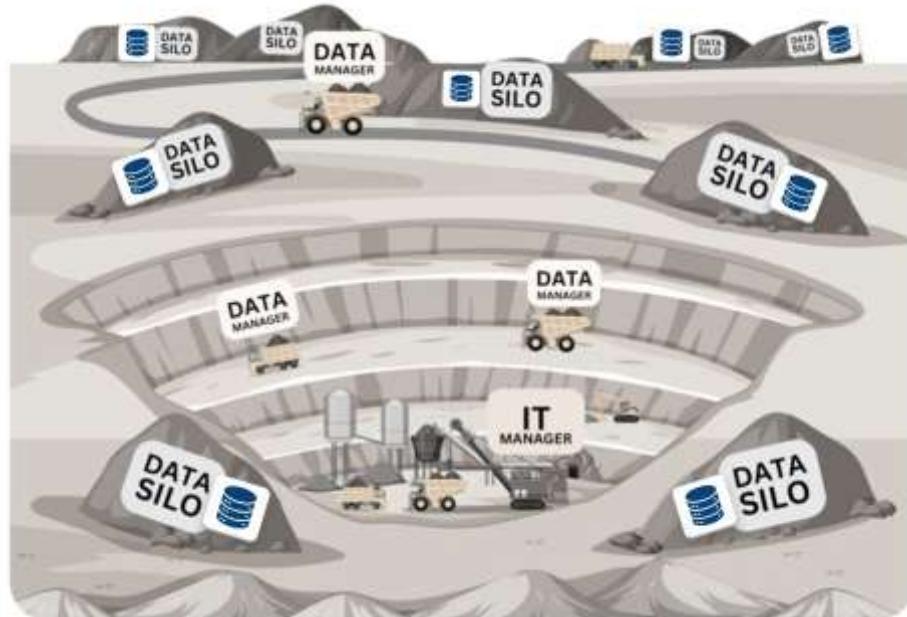
ir vērtīga informācija, bet to savstarpējā savienojuma trūkums neļauj izveidot vienotu, savstarpēji saistītu ekosistēmu. Šī nošķirtība kavē datu plūsmu un ierobežo organizācijas spēju redzēt pilnu ainu. Šo silosu savienošana ir ārkārtīgi ilgs un sarežģīts process, kurā vadības līmenī jāaudzē sēnu micēlijs, lai iemācītos, kā pārnest atsevišķus informācijas elementus starp sistēmām.

Saskaņā ar WEF 2016. gada pētījumu viens no galvenajiem digitālās transformācijas šķēršļiem ir vienotu datu standartu trūkums un sadrumstalotība.

*Būvniecības nozare ir viena no visvairāk sadrumstalotajām pasaulei, un tā ir atkarīga no visu vērtību kēdes dalībnieku vienmērīgas mijedarbības [5].*

- Pasaules ekonomikas forums 2016: nākotnes veidošana

Dizaineri, vadītāji, koordinatori un izstrādātāji bieži izvēlas strādāt autonomi, lai izvairītos no sarežģītas koordinācijas. Šī dabiskā tieksme novērtēta pie informācijas "silosu" izveides, kuros dati ir izolēti atsevišķās sistēmās. Jo vairāk ir šādu izolētu sistēmu, jo grūtāk ir panākt to sadarbību. Laika gaitā katrai sistēmai rodas sava datu bāze un specializēta vadītāju atbalsta nodaļa ( 1.2-4. attēls), kas vēl vairāk sarežģī integrāciju.



Attēls 2.1-5 Katrai sistēmai ir tendence radīt savu unikālu datu krātuvi, kas jāapstrādā ar piemērotiem rīkiem [31].

Korporatīvo sistēmu apburtais loks izskatās šādi: uzņēmumi iegulda līdzekļus sarežģītos izo leātos risinājumos, pēc tam saskaras ar augstām to integrācijas izmaksām, un izstrādātāji, apzinoties sistēmu apvienošanas sarežģību, izvēlas strādāt savās slēgtās ekosistēmās. Tas viss palielina IT ainavas sadrumstalotību un apgrūtina pāreju uz jauniem risinājumiem ( 2.1-5. attēls). Vadītāji galu galā kritizē datu silosus, bet reti analizē to cēloņus un to, kā tos novērst. Vadītāji sūdzas par novecojušām

IT sistēmām, bet to nomaiņa prasa ievērojamus ieguldījumus un reti dod gaidītos rezultātus. Rezultātā pat mēģinājumi apkarot problēmu bieži pasliktina situāciju.

Galvenais iemesls šādai atslēgšanai ir prioritātes piešķiršana lietojumprogrammām, nevis datiem. Uzņēmumi vispirms izstrādā atsevišķas sistēmas vai pērk gatavus risinājumus no piegādātājiem, un pēc tam mēģina tās apvienot, izveidojot dublējošas un nesaderīgas datu glabātuves un datubāzes.

Lai atrisinātu sadrumstalotības problēmu, ir nepieciešama radikāli jauna pieeja - prioritātes piešķiršana datiem, nevis lietojumprogrammām. Uzņēmumiem vispirms ir jāizstrādā datu pārvaldības stratēģijas un datu modeļi un pēc tam jāizveido sistēmas vai jāiegādājas risinājumi, kas strādā ar vienotu informācijas kopumu, nevis jārada jauni šķēršļi.

*Mēs ieejam jaunā pasaulei, kurā dati var būt svarīgāki par programmatūru.*

- Tims O'Reilijs, O'Reilly Media, Inc. izpilddirektors.

McKinsey Global Institute pētījums "Rethinking Construction: the path to improved productivity" (2016) liecina, ka būvniecības nozare atpaliek no citām nozarēm digitālās transformācijas jomā [32]. Saskaņā ar ziņojumu automatizētas datu pārvaldības un digitālo platformu ieviešana var ievērojami uzlabot produktivitāti un samazināt ar procesu nekonsekvenči saistītos zaudējumus. Šī digitālās transformācijas nepieciešamība ir uzsvērta arī Egana (Apvienotā Karaliste, 1998) ziņojumā [33], kurā uzsvērta integrētu procesu un sadarbības piejas galvenā nozīme būvniecībā.

Rezultātā, ja iepriekšējos 10 000 gados datu pārvaldītāju galvenā problēma bija datu trūkums, tad līdz ar datu un datu pārvaldības sistēmu lavīnu lietotāji un pārvaldītāji saskaras ar problēmu - datu pārbagātību, kas apgrūtina juridiski pareizas un kvalitatīvas informācijas atrašanu.

Atšķirīgi datu silosi neizbēgami novēd pie nopietnas datu kvalitātes samazināšanās problēmas. Vairāku neatkarīgu sistēmu gadījumā vieni un tie paši dati var pastāvēt dažādās versijās, bieži vien ar pretrunīgām vērtībām, radot papildu grūtības lietotājiem, kuriem jānosaka, kura informācija ir būtiska un uzticama.

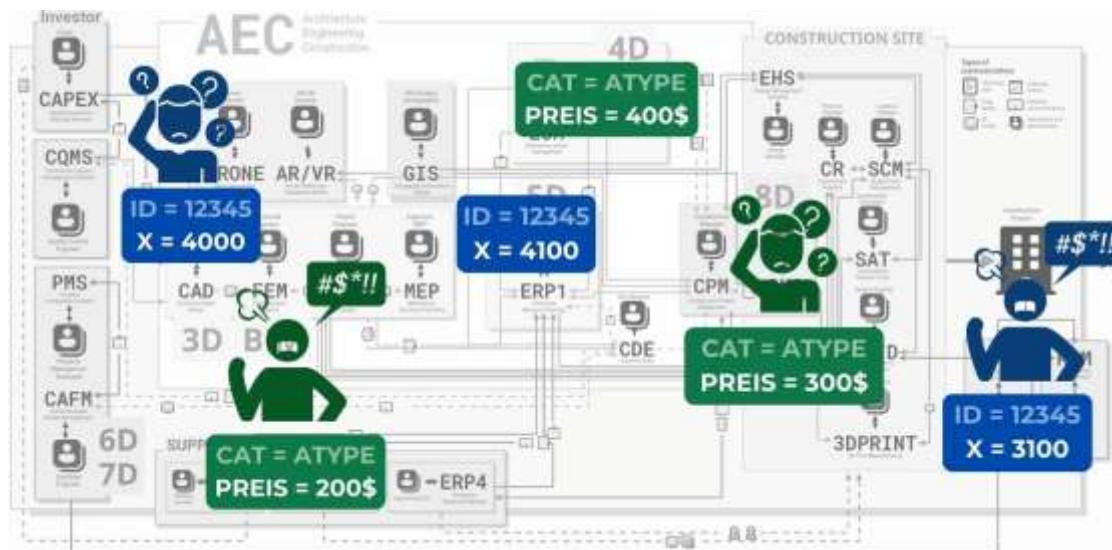
## Datu dublēšanās un datu kvalitātes trūkums, kas rodas nesaskaņotības dēļ.

Datu "silosu" problēmas dēļ vadītājiem ir jāpavada daudz laika datu meklēšanai un saskaņošanai. Lai nodrošinātos pret kvalitātes problēmām, uzņēmumi veido sarežģītas informācijas pārvaldības struktūras, kurās par datu meklēšanu, pārbaudi un saskaņošanu ir atbildīgi vairāki vadītāji. Tomēr šāda pieeja tikai palielina birokrātiju un palēnina lēmumu pieņemšanu. Jo vairāk ir datu, jo grūtāk tos analizēt un interpretēt, jo īpaši, ja nav vienota standarta to uzglabāšanai un apstrādei.

Tā kā pēdējā desmitgadē programmatūras lietojumprogrammu un sistēmu ir kļuvis daudz un tās aug-

kā sēnes pēc lietus, galalietotājiem aizvien aktuālāka ir kļuvusi "silos" un neatbilstošas datu kvalitātes problēma. Vienus un tos pašus datus, bet ar dažādām vērtībām, tagad var atrast dažādās sistēmās un lietojumprogrammās ( 2.1-6. attēls). Tas rada grūtības galalietotājiem, mēģinot noteikt, kura no daudzajām pieejamajām datu versijām ir atbilstoša un pareiza. Tas noved pie kļūdām analīzēs un galu galā lēmumu pieņemšanā.

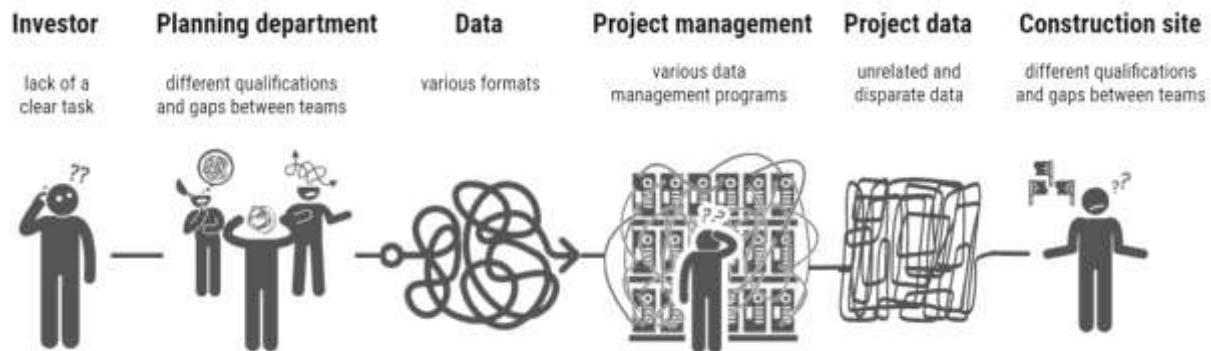
Lai apdrošinātos pret problēmām atrast pareizos datus, uzņēmuma vadītāji izveido daudzīmenū pārbaudes vadītāju birokrātiju. Viņu uzdevums ir spēt ātri atrast, pārbaudīt un nosūtīt vajadzīgos datus tabulu un pārskatu veidā, orientējoties atšķirīgu sistēmu labirintā.



**Attēls 2.1-6 Mēģinot atrast pareizos datus, vadītājiem ir jānodrošina datu kvalitāte un juridiskā uzticamība starp dažādām sistēmām.**

Tomēr praksē šis modelis rada jaunus sarežģījumus. Ja dati tiek pārvaldīti manuāli un informācija ir izkliedēta daudzos nesaistītos lēmumos, katrs mēģinājums iegūt precīzu un aktuālu informāciju, izmantojot lēmumu pieņēmēju piramīdu ( 2.1-7. attēls), klūst par šauru vietu - laikielpīgs un ar augstu kļūdu iespējamību.

Situāciju saasina digitālo risinājumu lavīna. Programmatūras tirgu turpina pārpludināt jauni, šķietami daudzsološi rīki. Taču bez skaidras datu pārvaldības stratēģijas šie risinājumi neintegrējas vienotā sistēmā, bet gan rada papildu sarežģītības un dublēšanās slāņus. Rezultātā tā vietā, lai vienkāršotu procesus, uzņēmumi nonāk vēl sadrumstalotākā un haotiskākā informācijas vidē.



**Attēlā: 2.1-7 Sistēmu sarežģītība un datu formātu daudzveidība rada konsekences zudumu būvniecības procesā.**

Visas šīs problēmas, kas saistītas ar daudzu atšķirīgu risinājumu pārvaldību, agri vai vēlu noved uzņēmuma vadību pie svarīgas atziņas: tas nav datu apjoms vai nākamā "universālā" datu apstrādes rīka meklēšana. Patiesais iemesls slēpjelas datu kvalitātē un tajā, kā organizācija tos rada, saņem, glabā un izmanto.

Ilgtspējīgu panākumu atslēga ir nevis jaunu "maģisku" lietojumprogrammu meklēšana, bet gan datu kultūras veidošana uzņēmumā. Tas nozīmē uzskatīt datus par stratēģisku aktīvu un noteikt datu kvalitāti, integritāti un atbilstību par prioritāti visos organizācijas līmenos.

Kvalitātes un kvantitātes dilemmas risinājums ir vienotas datu struktūras izveide, kas novērš dublēšanos, novērš nekonsekences un unificē informācijas plūsmas. Šī struktūra nodrošina vienotu, uzticamu datu avotu, uz kura pamata var pieņemt pamatotus, precīzus un savlaicīgus lēmumus.

Pretējā gadījumā, kā tas joprojām bieži notiek, uzņēmumi turpina paļauties uz HiPPO ekspertu subjektīviem viedokļiem un intuitīviem vērtējumiem, nevis uz ticamiem faktiem. Būvniecības nozarē, kur ekspertīzei tradicionāli ir nozīmīga loma, tas ir īpaši jūtams.

### **HiPPO jeb viedokļu bīstamība lēmumu pieņemšanā**

Tradicionāli būvniecības nozarē galvenie lēmumi tiek pieņemti, pamatojoties uz pieredzi un subjektīvu vērtējumu. Ja nav savlaicīgu un uzticamu datu, uzņēmumu vadītājiem ir jārīkojas akli, paļaujoties nevis uz objektīviem faktiem, bet gan uz vislabāk atalgotu darbinieku intuīciju (HiPPO - Highest Paid Person's Opinion) ( 2.1-8. attēls).

NO ANALYTICS?  
WELCOME TO THE HIPPO\*

\*HIGHEST PAID PERSON'S OPINION



**Attēls 2.1-8 Ja nav analītikas uzņēmējdarbība ir atkarīga no pieredzējušu speciālistu subjektīvā viedokļa.**

Šāda pieeja var būt attaisnojama stabilā un lēni mainīgā vidē, taču digitālās transformācijas laikmetā tā kļūst par nopietnu risku. Lēmumi, kas balstīti uz intuīciju un minējumiem, ir pakļauti izkroplojumiem, bieži balstās uz nepamatotām hipotēzēm un neņem vērā sarežģīto ainu, kas atspoguļojas datos

Tas, ko uzņēmumā lēmumu pieņemšanas līmenī uzskata par saprātīgām debatēm, bieži vien nav balstīts uz konkrētiem faktiem. Uzņēmuma panākumiem nevajadzētu būt atkarīgiem no ekspertu autoritātes un algas, bet gan no spējas efektīvi strādāt ar datiem, noteikt likumsakarības un pieņemt pamatotus lēmumus.

Ir svarīgi atteikties no priekšstata, ka autoritāte vai pieredze automātiski nozīmē, ka lēmums ir pareizs. Uz datiem balstīta pieeja maina spēles noteikumus: tagad lēmumu pieņemšanas pamatā ir dati un analītika, nevis amats un alga. Lielie dati, mašīnmācīšanās un vizuālā analītika ļauj mums noteikt likumsakarības un balstīties uz faktiem, nevis uz minējumiem ( 1.1-4. attēls).

*Bez datiem jūs esat tikai vēl viens cilvēks ar savu viedokli [34].*

- W. Edwards Deming, zinātnieks un vadības konsultants

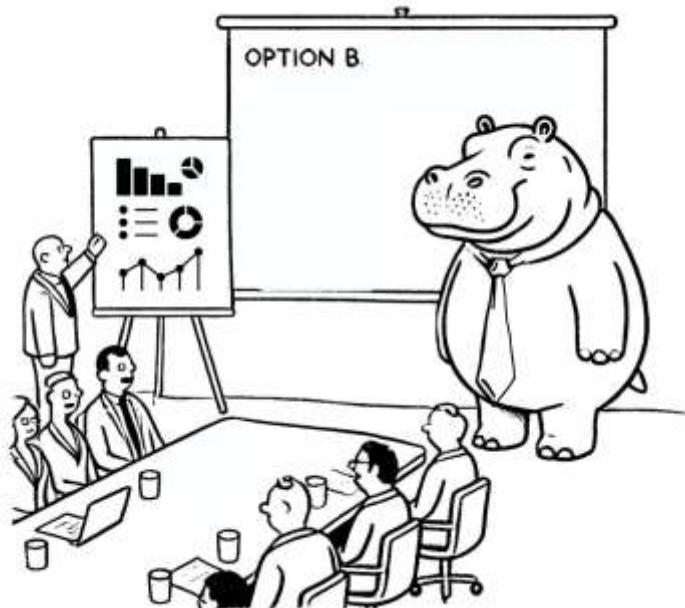
Modernas datu pārvaldības metodes nodrošina arī zināšanu nepārtrauktību uzņēmumā. Skaidri aprakstīti procesi, automatizācija un sistemātiska pieeja ļauj nodot pat galvenos amatus, nezaudējot efektivitāti.

Tomēr akla uzticēšanās datiem var novest arī pie nopietnām kļūdām. Datu paši par sevi ir tikai skaitļu kopums. Bez pienācīgas analīzes, konteksta un spējas identificēt likumsakarības tiem nav vērtības un tie nevar virzīt procesus. Panākumu atslēga slēpjās nevis izvēlē starp HiPPO intuīciju un analītiku, bet gan inteliģentu rīku izveidē, kas pārveido atšķirīgu informāciju pārvaldāmos, informētos lēmumos.

Digitālajā būvniecības vidē par izšķirošajiem veiksmes faktoriem kļūst nevis darba stāžs un vieta hierarhijā, bet gan spēja reaģēt, lēmumu precizitāte un resursu efektivitāte

Dati ir instrumenti, nevis absolūtas patiesības. Tiem ir jāpapildina cilvēka domāšana, nevis jāaizstāj tā. Neraugoties uz analītikas priekšrocībām, dati nevar pilnībā aizstāt cilvēka intuīciju un pieredzi. To uzdevums ir palīdzēt pieņemt precīzākus un pamatotākus lēmumus.

Konkurences priekšrocības tiks panāktas ne tikai ar standartu ievērošanu, bet arī ar spēju pārspēt konkurentus, efektīvi izmantojot resursus, kas visiem ir vienādi. Nākotnē datu prasmes kļūs tikpat svarīgas kā kādreiz rakstīprasme vai matemātikas prasmes. Speciālisti, kas spēj analizēt un interpretēt datus, spēs pieņemt precīzākus lēmumus, izspiežot tos, kas paļaujas tikai uz personīgo pieredzi ( 2.1-9. attēls).



**Attēls: 2.1-9 Lēmumiem jābalstās uz objektīvu analīzi, nevis uz augstāk atalgoтай darbinieka viedokli.**

Vadītāji, speciālisti un inženieri darbosies kā datu analītiķi, pētot projektu struktūru, dinamiku un galvenos rādītājus. Cilvēkresursi kļūs par sistēmas elementiem, kam būs nepieciešama elastīga, uz datiem balstīta pielāgošana, lai maksimāli palielinātu efektivitāti.

*Kļūdas, izmantojot nepietiekamus datus, ir daudz mazākas nekā tad, ja datu nav [35].*

- Čārlzs Bebidžs, pirmās analītiskās skaitļošanas mašīnas izgudrotājs

Lielo datu parādīšanās un LLM (Large Language Models) ieviešana ir radikāli mainījusi ne tikai

analīzes veidu, bet arī lēmumu pieņemšanas būtību. Ja agrāk galvenā uzmanība tika pievērsta cēloņsakarībām (kāpēc kaut kas ir noticis - diagnostiskā analīze) ( 1.1-4. attēls), tad šodien priekšplānā izvirzās spēja prognozēt nākotni (predikatīvā analīze) un nākotnē - preskriptīvā analīze, kad mašīnmācīšanās un mākslīgais intelekts iesaka labāko izvēli lēmumu pieņemšanas procesā.

Saskaņā ar jauno SAP™ pētījumu "New Study Finds Nearly Half of Executives Trust Artificial Intelligence More Than Themselves" 2025 [36], 44% augstākā līmeņa vadītāju būtu gatavi mainīt savu iepriekšējo lēmumu, pamatojoties uz mākslīgā intelekta padomu, un 38% uzticētos mākslīgajam intelektam pieņemt biznesa lēmumus savā vārdā. Tikmēr 74% vadītāju apgalvo, ka uzticas AI padomiem vairāk nekā saviem draugiem un ģimenei, un 55% strādā uzņēmumos, kuros ar AI iegūtas atziņas aizstāj vai bieži apiet tradicionālās lēmumu pieņemšanas metodes - īpaši organizācijās, kuru gada ieņēmumi pārsniedz 5 miljardus ASV dolāru. Turklāt 48% respondentu izmanto ģeneratīvos AI rīkus katru dienu, tostarp 15% tos izmanto vairākas reizes dienā.

Līdz ar LLM un automatizētu datu pārvaldības sistēmu attīstību rodas jauns izaicinājums: kā efektīvi izmantot informāciju, nezaudējot tās vērtību nesaderīgu formātu un neviendabīgu avotu haosā, ko papildina pieaugošā uzņēmējdarbības procesu sarežģītība un dinamika.

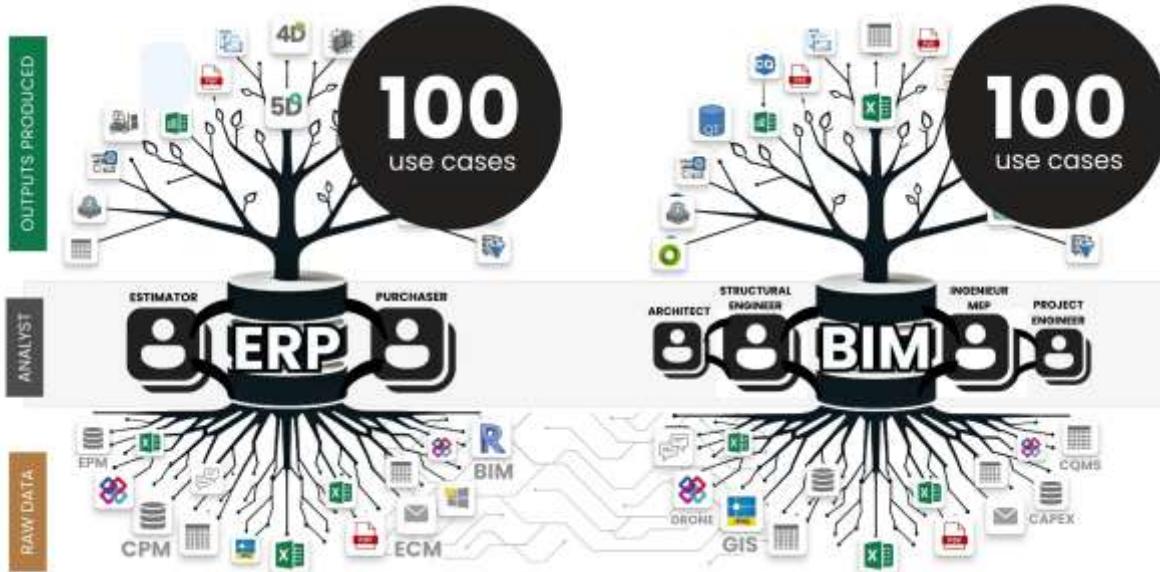
## **Nepārtraukti pieaug uzņēmējdarbības procesu sarežģītība un dinamika.**

Būvniecības nozare mūsdienās saskaras ar nopietniem izaicinājumiem datu un procesu pārvaldības jomā. Galvenās problēmas ir informācijas sistēmu atsveinotība, pārmērīga birokrātija un digitālo rīku integrācijas trūkums. Šīs problēmas klūst aizvien sarežģītākas, jo paši uzņēmējdarbības procesi klūst arvien sarežģītāki, ko nosaka tehnoloģijas, mainīgās klientu prasības un mainīgie noteikumi.

Būvprojektu unikalitāti nosaka ne tikai to tehniskās īpatnības, bet arī atšķirīgie nacionālie standarti un normatīvās prasības dažādās valstīs ( 4.2-10. attēls, 5.1-7. attēls). Tas prasa elastīgu, individuālu pieeju katram projektam, ko ir grūti īstenot tradicionālajās moduļu kontroles sistēmās. Procesu sarežģītības un lielā datu apjoma dēļ daudzi uzņēmumi vēršas pie piegādātājiem, kas piedāvā specializētus risinājumus. Taču tirgus ir pārslogots - daudzi jaunuzņēmumi piedāvā līdzīgus produktus, koncentrējoties uz šauriem uzdevumiem. Rezultātā bieži tiek zaudēta holistiska pieeja datu pārvaldībai.

Pielāgošanās nepārtrauktai jaunu tehnoloģiju plūsmai un tirgus prasībām klūst par izšķirošu konkurētspējas faktoru. Tomēr esošajām patentētajām lietojumprogrammām un modulārajām sistēmām ir zema pielāgošanās spēja - jebkuras izmaiņas bieži prasa ilgstošas un dārgas pārstrādes, ko veic izstrādātāji, kuri ne vienmēr izprot būvniecības procesu specifiku.

Uzņēmumi klūst par tehnoloģiskās atpalicības kīlniekiem, gaidot jaunus atjauninājumus, nevis nekavējoties ieviešot inovatīvas integrētas pieejas. Rezultātā būvniecības organizāciju iekšējā struktūra bieži vien ir sarežģīta savstarpēji saistītu hierarhisku un bieži vien slēgtu sistēmu ekosistēma, ko koordinē daudzīmeņu vadītāju tīkls ( 2.1-10. attēls).

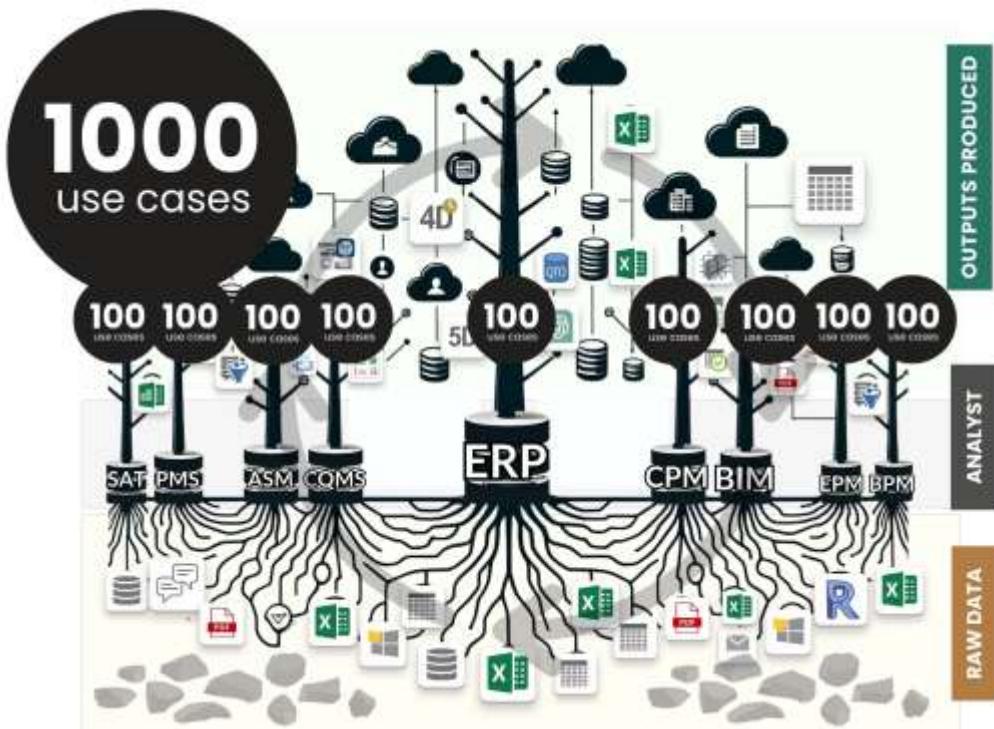


Attēls 2.1-10 Uzņēmumi sastāv no savstarpēji saistītām sistēmām, kuru savstarpējā savienošana veido procesus, kam nepieciešama automatizācija.

Saskaņā ar Kanādas Būvniecības asociācijas un KPMG Canada 2021. gadā veiktās aptaujas datiem [37] tikai 25% uzņēmumu uzskata, ka tehnoloģiju ieviešanas vai digitālo risinājumu ieviešanas jomā viņi atrodas nozīmīgā vai atšķirīgā pozīcijā salīdzinājumā ar konkurentiem. Tikai 23% respondentu norādīja, ka viņu risinājumi ir būtiski vai lielā mērā balstīti uz datiem. Tajā pašā laikā lielākā daļa respondentu virkni citu tehnoloģiju izmantošanu raksturoja kā tīri eksperimentālu vai atzina, ka tās vispār neizmanto.

Šī nevēlēšanās piedalīties tehnoloģiskos eksperimentos ir īpaši redzama lielos infrastruktūras projektos, kur kļūdas var izmaksāt miljoniem dolāru. Pat vismodernākās tehnoloģijas - digitālie dvīņi, prognozēšanas analītika - bieži sastopas ar pretestību nevis to efektivitātes, bet gan reālos projektos pierādītas uzticamības trūkuma dēļ.

Saskaņā ar Pasaules Ekonomikas foruma (WEF) ziņojumu "Veidojot būvniecības nākotni" [5], jauno tehnoloģiju ieviešana būvniecībā saskaras ne tikai ar tehniskām grūtībām, bet arī ar psiholoģiskiem šķēršļiem no klientu puses. [5], jaunu tehnoloģiju ieviešana būvniecībā saskaras ne tikai ar tehniskām grūtībām, bet arī ar psiholoģiskiem šķēršļiem no klientu puses. Daudzi klienti baidās, ka progresīvu risinājumu izmantošana padarīs viņu projektus par eksperimentālu objektu un padarīs viņus par "izmēģinājuma izmēģinājuma trusīšiem", un neparedzamas sekas var radīt papildu izmaksas un riskus.



Attēls 2.1-11 Katram datu izmantošanas gadījumam risinājumu tirgus piedāvā lietojumprogrammas procesu optimizācijai un automatizācijai.

Būvniecības nozare ir ļoti daudzveidīga: dažādiem projektiem ir atšķirīgas prasības, reģionālās īpatnības, normatīvie klasifikācijas noteikumi (4.2-10. attēls), aprēķinu standarti (5.1-7. attēls) utt. Tāpēc ir praktiski neiespējami izveidot patentētu universālu lietojumprogrammu vai sistēmu, kas perfekti atbilstu visām šīm prasībām un projektu specifikai.

Cenšoties tikt galā ar pieaugošo sistēmu sarežģītību un atkarību no programmatūras piegādātājiem, arvien vairāk tiek apzināts, ka efektīvas datu pārvaldības atslēga ir ne tikai atvērtība un standartizācija, bet arī pašas procesu arhitektūras vienkāršošana. Pieaugošā uzņēmējdarbības procesu sarežģītība un dinamika prasa jaunas pieejas, kurās prioritāte tiek pārcelta no datu uzkrāšanas uz to strukturēšanu un organizēšanu. Tieši šī pāreja būs nākamais solis būvniecības nozares attīstībā, kas iezīmēs programmatūras piegādātāju dominances ēras beigas un nozīmīgas informācijas organizācijas ēras sākumu.

Apzinoties universālu risinājumu ierobežojumus un neaizsargātību pret pieaugošo sarežģītību, prioritātes tiek mainītas no slēgtām platformām un datu uzkrāšanas uz pārredzamību, pielāgojamību un strukturētu informācijas apstrādi. Šī domāšanas maiņa atspoguļo plašākas pārmaiņas globālajā ekonomikā un tehnoloģijās, kas aprakstītas, izmantojot tā saukto "industriālo revolūciju" prizmu. Lai saprastu, kurp virzās būvniecība un kāds ir tās nākotnes virziens, ir jāapsver nozares vieta ceturtās un piektās industriālās revolūcijas kontekstā - no automatizācijas un digitalizācijas līdz personalizācijai, atvērtiem standartiem un uz pakalpojumiem balstītam datu modelim.

## Ceturtā industriālā revolūcija (Industry 4.0) un piektā industriālā revolūcija (Industry 5.0) būvniecībā.

Tehnoloģiskie un ekonomiskie posmi ir teorētiski jēdzieni, ko izmanto, lai aprakstītu un analizētu sabiedrības un ekonomikas attīstību dažādos attīstības posmos. Dažādi pētnieki un eksperti tos var interpretēt dažādi.

- **Ceturtā industriālā revolūcija** (4IR jeb Industry 4.0) ir saistīta ar informācijas tehnoloģijām, automatizāciju, digitalizāciju un globalizāciju. Viens no tās galvenajiem elementiem ir patentētu programmatūras risinājumu, t. i., specializētu digitālo produktu, kas paredzēti konkrētiem uzdevumiem un uzņēmumiem, izveide. Šie risinājumi bieži vien kļūst par nozīmīgu IT infrastruktūras daļu, bet bez papildu modifikācijām ir slikti mērogojami.
- **Piektā industriālā revolūcija** (5IR) šobrīd ir agrīnā konceptualizācijas un attīstības stadijā nekā 4IR. Tās pamatprincipi ietver lielāku produktu un pakalpojumu personalizāciju. 5IR ir virzība uz pielāgojamāku, elastīgāku un personalizētāku saimniecisko darbību, koncentrējoties uz personalizāciju, konsultācijām un uz pakalpojumiem orientētiem modeļiem. Piektā ekonomikas veida galvenais aspekts ir datu izmantošana lēmumu pieņemšanā, kas praktiski nav iespējama bez atvērto datu un atvērto rīku izmantošanas (2.1-12. attēls).

FOURTH INDUSTRIAL REVOLUTION  
(4IR OR INDUSTRY 4.0)



FIFTH INDUSTRIAL REVOLUTION  
(5IR)



Attēls 2.1-12 Ceturtais modelis ir vērts uz risinājumiem, savukārt piektais modelis ir vērts uz personalizāciju un datiem.

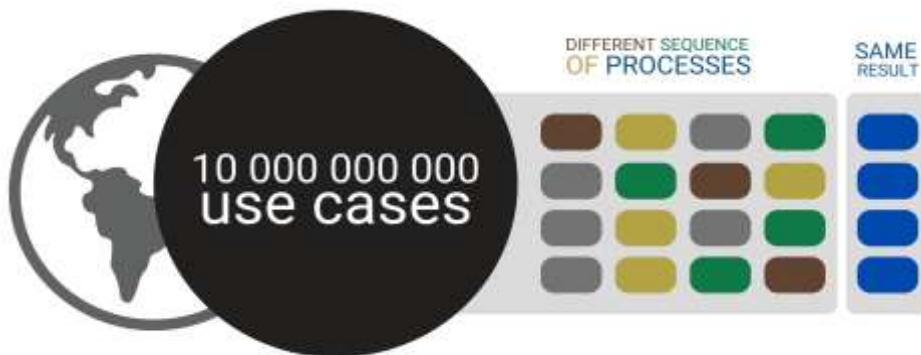
Būvniecības nozares uzņēmumiem paredzētas lietojumprogrammas izveide izmantošanai desmit vai simts organizācijās negarantē tās veiksmīgu izplatīšanu citos uzņēmumos, reģionos vai valstīs bez būtiskiem pārveidojumiem un uzlabojumiem. Šādu risinājumu veiksmīgas mērogošanas varbūtība joprojām ir neliela, jo katrai organizācijai ir unikāli procesi, prasības un apstākļi, kas var prasīt individuālus pielāgojumus.

Ir svarīgi saprast, ka jau šodien veiksmīga tehnoloģisko risinājumu integrācija nozīmē individuālu pieeju katram procesam, projektam un uzņēmumam. Tas nozīmē, ka pat pēc tam, kad ir izstrādāta universāla sistēma, rīks vai programma, tā būs detalizēti jāpielāgo un jāpielāgo, lai atbilstu katra konkrētā uzņēmuma un projekta unikālajām prasībām un nosacījumiem.

Saskaņā ar PwC ziņojumu "Decoding the Fifth Industrial Revolution" [38]. [38] [38], šogad aptuveni 50% dažādu nozaru augstāko vadītāju paļaujas uz progresīvo tehnoloģiju un cilvēku zināšanu integrāciju. Šāda pieeja ļauj ātri pielāgoties izmaiņām produktu dizainā vai klientu prasībās, radot personalizētu ražošanu.

Katram procesam ir jāizstrādā unikāla funkcija vai lietojumprogramma, kas, nemit vērā pasaules būvniecības nozares lielumu un projektu daudzveidību, noved pie tā, ka pastāv milzīgs skaits uzņēmējdarbības gadījumu, kas katru reizi pārstāv unikālu cauruļvadu logiku ( 2.1-13. attēls). Katram šādam gadījumam ir savas īpatnības, un tam nepieciešama pielāgota pieeja. Mēs sīkāk aplūkosim vienas un tās pašas analītiskās problēmas iespējamo risinājumu daudzveidību dažādu pieeju kontekstā nodaļā, kas veltīta mašīnmācībai un "Titānika" datu kopas analīzei ( 9.2-9. attēls).

**Cauruļvads** digitālo procesu kontekstā ir darbību, procesu un rīku secība, kas nodrošina automatizētu vai strukturētu datu un darba plūsmu dažādos projekta dzīves cikla posmos.



Attēls 2.1-13 Biznesa gadījumu individualitāte un mainīgums padara neiespējamus mēģinājumus izveidot mērogojamas slēgtas platformas un rīkus.

Digitalās transformācijas ietekmē mūsu dzīve jau ir daudzējādā ziņā mainījusies, un šodien mēs varam runāt par jauna būvniecības nozares ekonomiskās attīstības posma sākumu. Šajā "jaunajā ekonomikā" konkurence tiks organizēta pēc citiem noteikumiem: tas, kurš spēs efektīvi pārvērst publiskās zināšanas un atklātos datus pieprasītos produktos un pakalpojumos, iegūs galveno priekšrocību piektās industriālās revolūcijas apstākļos.

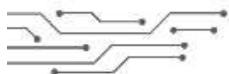
Kā norāda ekonomiste Kate Maskus grāmatā "Privātās tiesības un sabiedriskās problēmas: Globālā intelektuālā īpašuma ekonomika 21. gadsimtā" [39]. 2012 [39], "mēs dzīvojam globālā zināšanu ekonomikā, un nākotne pieder tiem, kas zina, kā zinātniskos atklājumus pārvērst precēs".

Pāreja uz piekto ekonomikas režīmu nozīmē pāreju no slēgtiem IT risinājumiem uz atvērtiem standartiem un platformām. Uzņēmumi sāks atteikties no tradicionālajiem programmatūras produktiem par labu uz pakalpojumiem orientētiem modeļiem, kuru galvenais aktīvs būs dati, nevis patentētas tehnoloģijas.

Hārvarda Biznesa skolas 2024. gada pētījums [40] liecina par milzīgo ekonomisko vērtību, ko rada atvērtā koda programmatūra (Open Source Software, OSS). Saskaņā ar pētījuma datiem OSS ir 96% no visiem programmatūras kodiem, un dažas komerciālās programmatūras 99,9% sastāv no OSS komponentiem. Bez OSS uzņēmumi programmatūrai tērētu 3,5 reizes vairāk.

Uzņēmumu ekosistēmu veidošana, sekojot globālajām tendencēm, pakāpeniski pāries uz piekto ekonomisko paradigmu, kurā uz datiem orientēta analītika un konsultāciju pakalpojumi klūs par augstāku prioritāti nekā izolēti, slēgti risinājumi ar stingri definētiem izmantošanas scenārijiem.

Digitalizācijas laikmets mainīs spēku līdzsvaru nozarē: tā vietā, lai paļautos uz piegādātāju risinājumiem, uzņēmumi savu konkurētspēju balstīs uz spēju efektīvi izmantot datus. Tā rezultātā būvniecības nozare pāries no mantotām, neelastīgām sistēmām uz elastīgām, pielāgojamām ekosistēmām, kurās projektu pārvaldības pamatā būs atvērti standarti un savstarpēji savietojami rīki. Lietojumprogrammu piegādātāju dominances laikmeta beigas radīs jaunu vidi, kurā vērtību noteiks nevis slēgta pirmkoda un specializētu savienotāju turēšana, bet gan spēja pārvērst datus stratēģiskā priekšrocībā.



## NODAĻA 2.2.

### HAOSA PĀRVĒRŠANA KĀRTĪBĀ UN SAREŽGĪTĪBAS SAMAZINĀŠANA

#### Liekais kods un slēgtas sistēmas kā šķērslis produktivitātes uzlabošanai

Pēdējās desmitgadēs tehnoloģiskās pārmaiņas IT jomā galvenokārt ir veicinājuši programmatūras ražotāji. Viņi noteica attīstības virzienu, nosakot, kuras tehnoloģijas uzņēmumiem jāievieš un kuras jāatstāj novārtā. Laiķmetā, kad no izolētiem risinājumiem pārgāja uz centralizētām datubāzēm un integrētām sistēmām, pārdevēji popularizēja licencētus produktus, nodrošinot piekļuves un mērogojamības kontroli. Vēlāk, līdz ar mākoņtehnoloģiju un programmatūras kā pakalpojuma (SaaS) modeļu parādīšanos, šī kontrole pārtapa par abonēšanas modeli, nostiprinot lietotājus kā lojālus digitālo pakalpojumu klientus.

Šāda pieeja ir radījusi paradoksu: neskatoties uz vēl nepieredzēti lielo radīto programmu kodu apjomu, tikai neliela daļa no tā tiek faktiski izmantota. Iespējams, ir simtiem vai tūkstošiem reižu vairāk koda, nekā nepieciešams, jo vieni un tie paši uzņēmējdarbības procesi tiek aprakstīti un dublēti desmitiem vai simtiem programmu dažādos veidos pat viena uzņēmuma ietvaros. Tajā pašā laikā izstrādes izmaksas jau ir samaksātas, un šīs izmaksas ir neatgūstamas. Neraugoties uz to, nozare turpina šo ciklu, radot jaunus produktus ar minimālu pievienoto vērtību galalietotājam, biežāk tirgus gaidu, nevis reālo vajadzību spiediena ietekmē.

Saskaņā ar Aizsardzības iegādes universitātes (Defence Acquisition University (DAU) Software Development Cost Estimating Guide [41], programmatūras izstrādes izmaksas var ievērojami atšķirties atkarībā no vairākiem faktoriem, tostarp sistēmas sarežgītības un izvēlētās tehnoloģijas. Vēsturiski 2008. gada izstrādes izmaksas ir bijušas aptuveni 100 ASV dolāru par vienu pirmkoda rindu (SLOC), savukārt uzturēšanas izmaksas var sasniegt 4000 ASV dolāru par SLOC.

Tikai viena no CAD lietojumprogrammu sastāvdalām - ģeometriskais kodols - var saturēt desmitiem miljonu kodu rindu (6.1-5. attēls). Līdzīga situācija ir vērojama ERP sistēmās (5.4-4. attēls), pie kuru sarežgītības apspriešanas mēs atgriezīsimies grāmatas piektajā daļā. Tomēr, rūpīgāk ieskatoties, atklājas, ka liela daļa no šī koda nerada pievienoto vērtību, bet tikai darbojas kā "pastnieks" - mehāniski pārvieto datus starp datu bāzi, API, lietotāja saskarni un citām sistēmas tabulām. Neraugoties uz populāro mītu par tā sauktās biznesa logikas izšķirošo nozīmi, skarbā realitāte ir daudz prozaiskāka: mūsdienu kodu bāzes ir pilnas ar novecojušiem šablonu blokiem (legacy code), kuru vienīgais mērķis ir nodrošināt datu pārsūtīšanu starp tabulām un komponentēm, neietekmējot lēmumu pieņemšanu vai biznesa efektivitāti.

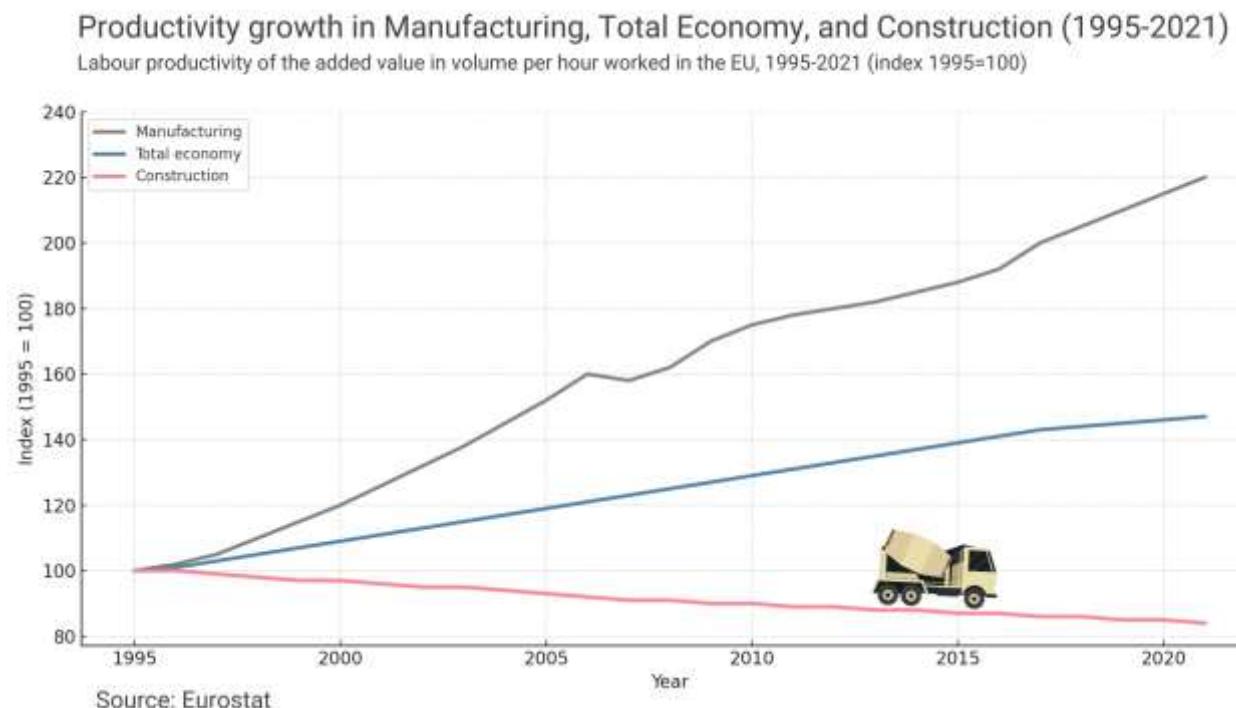
Tā rezultātā slēgti risinājumi, kas apstrādā datus no dažādiem avotiem, neizbēgami pārvēršas par jucekļīgām "spageti ekosistēmām". Šādas sarežgītas, savstarpēji saistītas sistēmas var apstrādāt tikai armija vadītāju, kas strādā daļēji rutīnas režīmā. Šāda datu pārvaldības organizācija ir ne tikai neefektīva resursu ziņā, bet arī rada kritiskas ievainojamības uzņēmējdarbības procesos, padarot uzņēmumu atkarīgu no šaura speciālistu loka, kuri izprot, kā šis tehnoloģiskais labirints darbojas.

Nepārtrauktais koda apjoma pieaugums, lietojumprogrammu skaita palielināšanās un arvien

sarežģītākās piegādātāju piedāvātās koncepcijas ir novedušas pie likumsakarīga rezultāta - IT ekosistēmas sarežģības pieauguma būvniecības nozarē. Tas ir padarījis neefektīvu digitalizācijas praktisko īstenošanu, palielinot lietojumprogrammu skaitu nozarē. Programmatūras produkti, kas radīti, nepievēršot pienācīgu uzmanību lietotāju vajadzībām, bieži vien prasa ievērojamus resursus ieviešanai un atbalstam, bet nedod gaidīto atdevi.

Saskaņā ar McKinsey pētījumu "Celtniecības produktivitātes palielināšana". [42], pēdējās divās desmitgadēs darba ražīguma pieaugums būvniecības nozarē pasaulei vidēji ir bijis tikai 1% gadā, salīdzinot ar 2,8% pieaugumu pasaules ekonomikā kopumā un 3,6% pieaugumu rūpniecībā. Amerikas Savienotajās Valstīs darba ražīgums būvniecībā uz vienu strādājošo kopš pagājušā gadsimta 60. gadiem ir samazinājies uz pusi [43].

Sistēmu sarežģības pieaugums, izolētība un slēgtie dati ir pasliktinājuši saziņu starp speciālistiem, padarot būvniecības nozari par vienu no vismazāk efektīvajām ( 2.2-1. attēls). Līdz 2040. gadam līdz 22 trilioniem ASV dolāru, kas prasīs ievērojamu efektivitātes pieaugumu.



**Attēlā: 2.2.2-1 Slēgti un sarežģīti dati, kā rezultātā vāja saziņa starp speciālistiem noveda būvniecības nozari pie vienas no visneefektīvākajām tautsaimniecības nozarēm (pēc [44], [45]).**

Kā uzsvērts McKinsey (2024) pētījumā "Būvniecības produktivitātes nodrošināšana vairs nav izvēles iespēja", ņemot vērā pieaugošo resursu nepietiekamību un nozares ambīcijas divkāršot izaugsmes tempu, būvniecība vairs nevar atļauties saglabāt pašreizējo produktivitātes līmeni [44]. Tieki prognozēts, ka globālās būvniecības izmaksas pieaugs no 13 trilioniem ASV dolāru 2023. gadā līdz daudz lielākam līmenim desmitgades beigās, tāpēc efektivitātes jautājums kļūst ne tikai aktuāls, bet kritiski svarīgs.

Viens no galvenajiem efektivitātes uzlabošanas veidiem būs neizbēgama lietojumprogrammu

struktūru un datu ekosistēmas arhitektūru unifikācija un vienkāršošana. Šāda racionalizācijas pieeja novērsīs liekos abstrakcijas slāņus un nevajadzīgu sarežģītību, kas gadu gaitā uzkrājušies uzņēmumu sistēmās.

## No atsevišķam datu noliktavām uz vienotu datu noliktavu

Jo vairāk datu organizācija uzkrāj, jo grūtāk no tiem iegūt reālu vērtību. Tā kā informācija tiek uzglabāta izolētos silosos, mūsdienu uzņēmumu uzņēmējdarbības procesi ir līdzīgi celtniekiem, kas cenšas uzbūvēt debesskrāpi no materiāliem, kas glabājas tūkstošiem dažādu noliktavu. Informācijas pārbagātība ne tikai apgrūtina piekļuvi juridiski svarīgai informācijai, bet arī palēnina lēmumu pieņemšanu: katrs solis ir atkārtoti jāpārbauda un jāapstiprina.

Katrs uzdevums vai process ir piesaistīts atsevišķai tabulai vai datubāzei, un datu apmaiņai starp sistēmām ir nepieciešama sarežģīta integrācija. Klūdas un neatbilstības vienā sistēmā var izraisīt kēdes kļūmes citās sistēmās. Nepareizas vērtības, novēloti atjauninājumi un dublējoša informācija liek darbiniekiem tērēt daudz laika manuālai datu saskaņošanai un salīdzināšanai. Rezultātā organizācija vairāk laika tērē sadrumstalotības seku novēršanai, nevis procesu izstrādei un optimizācijai

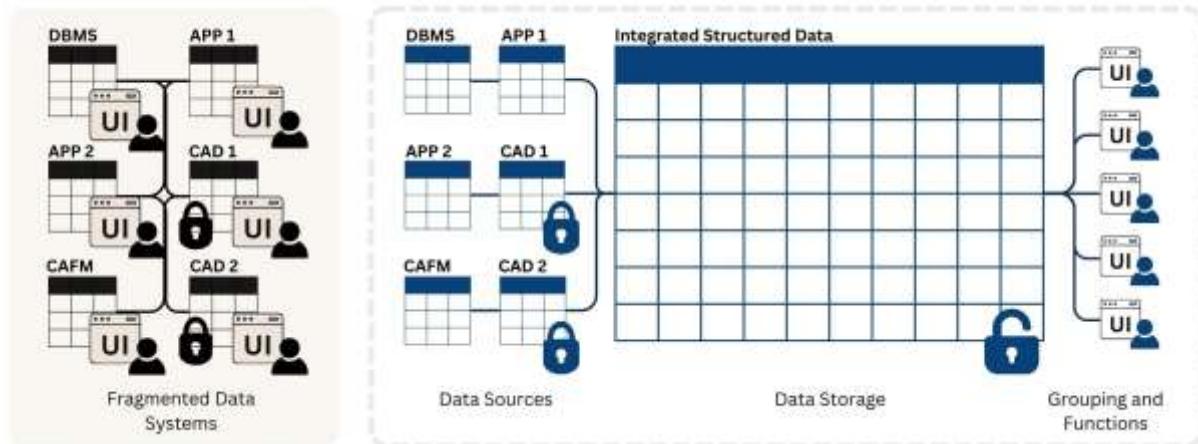
Šī problēma ir universāla: daži uzņēmumi turpina cīnīties ar haosu, bet citi atrod risinājumu integrācijā - informācijas plūsmu pārvietošanā uz centralizētu glabāšanas sistēmu. Iedomājieties to kā vienu lielu tabulu, kurā var uzglabāt visas ar uzdevumiem, projektiem un objektiem saistītās vienības. Desmitiem atšķirīgu tabulu un formātu vietā veidojas vienota, vienota, saskanīga krātuve (2.2.2-2. attēls), kas ļauj:

- samazināt datu zudumu;
- novērstu vajadzību pēc pastāvīgas informācijas saskaņošanas;
- uzlabot datu pieejamību un kvalitāti;
- vienkāršot analītisko apstrādi un mašīnmācīšanos.

Datu saskaņošana ar vienotu standartu nozīmē, ka neatkarīgi no avota informācija tiek pārveidota vienotā un mašīnlasāmā formātā. Šāda datu organizācija ļauj pārbaudīt to integritāti, analizēt tos reālajā laikā un nekavējoties izmantot vadības lēmumu pieņemšanai.

Integrēto glabāšanas sistēmu koncepcija un to pielietojums analīzē un mašīnmācībā tiks sīkāk aplūkota nodaļā "Lielo datu glabāšana un mašīnmācīšanās". Datu modelēšanas un strukturēšanas tēmas tiks detalizēti aplūkotas nodaļās "Datu pārveidošana strukturētā formā" un "Kā standarti maina

spēli: no nejaušiem failiem līdz izstrādātam datu modelim".



**Attēls 2.2-2 Datu integrācija novērš "silosus", uzlabo informācijas pieejamību un optimizē uzņēmējdarbības procesus.**

Kad dati ir strukturēti un apvienoti, nākamas loģiskais solis ir to validēšana. Izmantojot vienotu integrētu repozitoriju, šis process ir ievērojami vienkāršots: vairs nav vairāku nekonsekventu shēmu, dublētu struktūru un sarežģītu attiecību starp tabulām. Visa informācija ir saskaņota ar vienotu datu modeli, novēršot iekšējās nekonsekvences un paātrinot validēšanas procesu. Validēšana un datu kvalitātes nodrošināšana ir visu uzņēmējdarbības procesu stūrakmeņi, un mēs tos sīkāk aplūkosim attiecīgajās grāmatas nodalās.

Pēdējā posmā datus sagrupē, filtrē un analizē. Tiem tiek piemērotas dažādas funkcijas: apkopošana (saskaitīšana, reizināšana), aprēķini starp tabulām, kolonnām vai rindām ( 2.2-4. attēls). Darbs ar datiem kļūst par secīgu darbību: datu vākšana, strukturēšana, validēšana, transformēšana, analītiskā apstrāde un novirzīšana uz galīgajām lietojumprogrammām, kur informāciju izmanto praktisku problēmu risināšanai. Plašāk par šādu scenāriju veidošanu, soļu automatizēšanu un apstrādes plūsmu veidošanu mēs runāsim nodalās par ETL -procesiem un datu plūsmas pieeju.

Tādējādi digitālā transformācija nav tikai informācijas apstrādes vienkāršošana. Runa ir par pārmērīgas sarežģītības novēršanu datu pārvaldībā, pārejot no haosa uz prognozējamību, no daudzām sistēmām uz pārvaldāmu procesu. Jo mazāka ir arhitektūras sarežģītība, jo mazāk koda ir nepieciešams tās atbalstam. Nākotnē kods kā tāds var izzust pavisam, dodot vietu inteliģentiem aģentiem, kas patstāvīgi analizē, sistematizē un pārveido datus.

## **Integrētas glabāšanas sistēmas ļauj pāriet uz mākslīgā intelekta aģentiem**

Jo mazāk sarežģīti ir dati un sistēmas, jo mazāk koda ir jāraksta un jāuztura. Un visvienkāršākais veids, kā ietaupīt izstrādi, ir pilnībā atbrīvoties no koda, aizstājot to ar datiem. Kad lietojumprogrammu koda izstrāde pāriet no koda uz datu modeļiem, neizbēgami notiek pāreja uz uz datiem orientētu (uz datiem balstītu) pieeju, jo aiz šiem jēdzieniem slēpjās pavisam cits domāšanas veids.

Kad cilvēks izvēlas strādāt ar datiem centrā, viņš sāk to lomu uztvert citādi. Dati vairs nav tikai "izejmateriāls" lietojumprogrammām - tagad tie ir pamats, uz kura tiek veidota arhitektūra, loģika un mijiedarbība.

Tradicionālā pieeja datu pārvaldībai parasti sākas lietojumprogrammu līmenī un būvniecībā atgādina apgrūtinošu birokrātisku sistēmu: daudzīmenē apstiprinājumi, manuālas pārbaudes, bezgalīgas dokumentu versijas, izmantojot attiecīgos programmatūras produktus. Attīstoties digitālajām tehnoloģijām, arvien vairāk uzņēmumu būs spiesti pāriet uz minimālisma principu - uzglabāt un izmantot tikai to, kas patiešām ir nepieciešams un tiks izmantots.

Pārdevēji ir pārņēmuši minimizācijas loģiku. Lai vienkāršotu datu glabāšanu un apstrādi, lietotāju darbs tiek pārcelts no bezsaistes lietojumprogrammām un rīkiem uz mākoņpakaļpojumiem un tā sauktajiem SaaS risinājumiem.

SaaS koncepcija (Software as a Service jeb "programmatūra kā pakalpojums") ir viena no galvenajām mūsdienu IT infrastruktūras tendencēm, kas ļauj lietotājiem piekļūt lietojumprogrammām, izmantojot internetu, bez nepieciešamības instalēt un uzturēt programmatūru savā datorā.

No vienas puses, SaaS ir veicinājusi mērogošanu, versiju kontroli un samazinājusi atbalsta un uzturēšanas izmaksas, taču, no otras puses, papildus atkarībai no konkrētās lietojumprogrammas loģikas tā ir padarījusi lietotāju pilnībā atkarīgu no pakalpojumu sniedzēja mākoņa infrastruktūras. Ja pakalpojums nedarbojas, piekļuve datiem un uzņēmējdarbības procesiem var tikt uz laiku vai pat pastāvīgi bloķēta. Turklāt visi lietotāja dati, strādājot ar SaaS lietojumprogrammām, tiek glabāti pakalpojumu sniedzēja serveros, kas rada drošības un normatīvās atbilstības riskus. Tarifu vai lietošanas noteikumu izmaiņas var arī palielināt izmaksas vai radīt nepieciešamību steidzami migrēt.

Mākslīgā intelekta, LLM -aģentu un uz datiem orientētas pieejas attīstība ir apšaubījusi lietojumprogrammu nākotni to tradicionālajā formā un SaaS izpildi. Ja iepriekš lietojumprogrammām un pakalpojumiem bija jāpārvalda biznesa loģika un jāapstrādā dati, tad līdz ar mākslīgā intelekta aģentu parādīšanos šīs funkcijas var pāriet uz viedajām sistēmām, kas strādā tieši ar datiem.

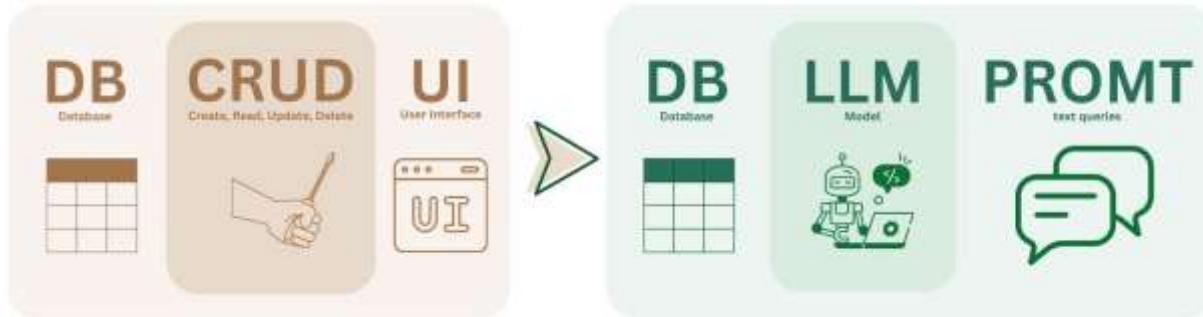
Tāpēc IT nodalās un vadības līmenī arvien biežāk tiek apspriestas hibrīdās arhitektūras, kurās mākslīgā intelekta -aģenti un lokālie risinājumi papildina mākoņpakaļpojumus, samazinot atkarību no SaaS -platformām.

*Mūsu pieeja atzīst, ka tradicionālās biznesa lietojumprogrammas vai SaaS lietojumprogrammas aģentu laikmetā var būtiski mainīties. Šīs lietojumprogrammas būtībā ir CRUD [izveidot, nolasīt, atjaunināt un dzēst] datubāzes ar biznesa loģiku. Taču nākotnē šo loģiku pārņems mākslīgā intelekta aģenti [46].*

- Satja Nadella, Microsoft izpilddirektors, 2024. gads.

Uz datiem orientēta pieeja un mākslīgā intelekta/LLM aģentu izmantošana var samazināt lieko procesu skaitu un tādējādi samazināt darbinieku darba slodzi. Ja dati ir pareizi organizēti, tos ir vieglāk analizēt, vizualizēt un izmantot lēmumu pieņemšanā. Bezgalīgu ziņojumu un pārbaužu vietā speciālisti ar dažiem klikšķiem vai ar LLM aģentu palīdzību automātiski iegūst piekļuvi aktuālajai informācijai gatavu dokumentu un paneļu veidā.

Datu apstrādē mums palīdzēs mākslīgā intelekta rīki (AI) un LLM tērzēšana. Pēdējos gados ir vērojama tendence datu pārvaldībā izmantot nevis tradicionālās CRUD operācijas (izveidot, nolasīt, atjaunināt, dzēst), bet gan lielos valodas modeļus (LLM). LLM spēj interpretēt dabisko valodu un automātiski ģenerēt atbilstošus datubāzes pieprasījumus, kas vienkāršo mijiedarbību ar datu pārvaldības sistēmām ( 2.2-3. attēls).



**Attēls 2.2-3 Mākslīgais intelekts aizstās un integrēs glabāšanas un datubāzu risinājumus, pakāpeniski izspiežot tradicionālās lietojumprogrammas un CRUD -operācijas.**

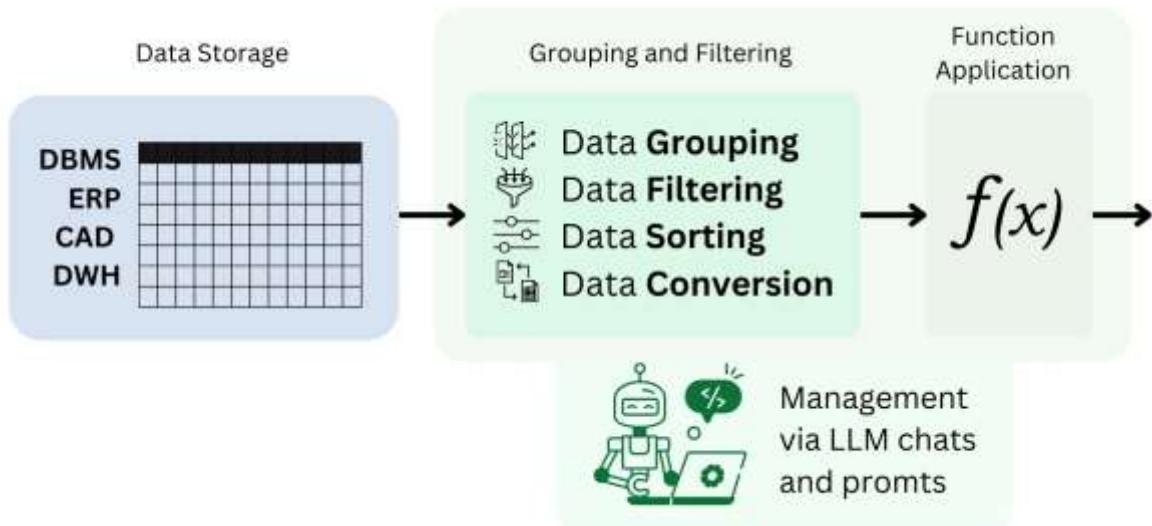
Nākamo 3-6 mēnešu laikā mākslīgais intelekts rakstīs 90% koda, bet 12 mēnešu laikā gandrīz visu kodu varētu ģenerēt mākslīgais intelekts [47].

- Dario Amodei, LLM izpilddirektors Anthropic, 2025. gada marts.

Neraugoties uz mākslīgā intelekta izstrādes rīku straujo attīstību (piemēram, GitHub Copilot), 2025. gadā izstrādātājiem joprojām būs galvenā loma šajā procesā. Mākslīgā intelekta aģenti kļūst par arvien nodeīgākiem palīgiem: tie automātiski interpretē lietotāja vaicājumus, ģenerē SQL un Pandas vaicājumus (vairāk par to nākamajās nodalās) vai raksta kodu datu analīzei. Šādā veidā mākslīgais intelekts pakāpeniski aizstāj tradicionālās lietojumprogrammu lietotāja saskarnes.

Mākslīgā intelekta modeļu, piemēram, valodas modeļu, izplatīšanās veicinās hibrīdās arhitektūras attīstību. Tā vietā, lai pilnībā atteiktos no mākoņrisinājumiem un SaaS produktiem, mēs varam redzēt mākoņpalkojumu integrāciju ar vietējām datu pārvaldības sistēmām. Piemēram, federatīvā mācīšanās īauj izveidot jaudīgus mākslīgā intelekta modeļus, nepārvietojot sensitīvus datus uz mākonī. Šādā veidā uzņēmumi var saglabāt kontroli pār saviem datiem, vienlaikus iegūstot piekļuvi

modernām tehnoloģijām.



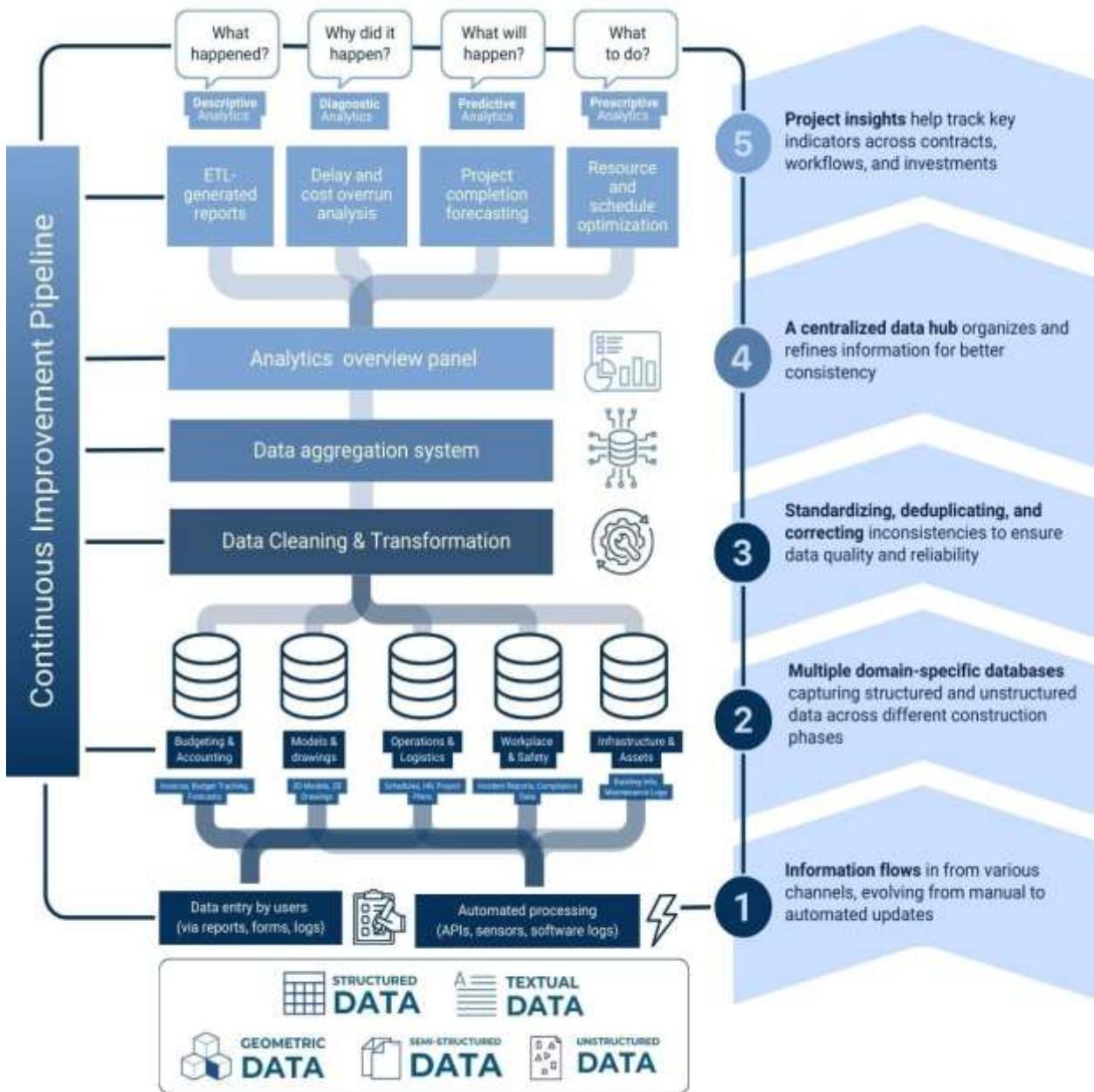
Attēls 2.2-4 Pamatoperācijas - grupēšana, filtrēšana un šķirošana, kam seko funkciju piemērošana - tiks veiktas ar LLM čatu.

Būvniecības nozares nākotne balstīsies uz lokālo risinājumu, mākoņa jaudas un inteliģento modeļu kombināciju, kas sadarbojas, lai izveidotu efektīvas un drošas datu pārvaldības sistēmas. LLM ļaus lietotājiem bez dzīlām tehniskām zināšanām mijiedarboties ar datu bāzēm un datu noliktavām, formulējot savus pieprasījumus dabiskajā valodā. Plašāk par LLM un mākslīgā intelekta aģentiem un to darbību mēs runāsim nodaļā "LLM aģenti un strukturēto datu formāti".

Pareizi organizēti dati un vienkārši, viegli lietojami analītiskie rīki, kas nodrošina LLM, ne tikai atvieglos darbu ar informāciju, bet arī palīdzēs līdz minimumam samazināt kļudas, palielināt efektivitāti un automatizēt procesus.

## No datu vākšanas līdz lēmumu pieņemšanai: ceļš uz automatizāciju

Turpmākajās grāmatas daļās mēs detalizēti aplūkosim, kā speciālisti mijiedarbojas savā starpā un kā dati kļūst par pamatu lēmumu pieņemšanai, automatizācijai un darbības efektivitātei. Attēlā 2.2.2-5 ir sniegs diagrammas piemērs, kurā parādīta datu apstrādes posmu secība datu centrētā pieejā. Šī diagramma ilustrē nepārtrauktas uzlabošanas cauruļvadu), kura daļas tiks detalizēti aplūkotas turpmāk grāmatā.



Attēls 2.2-5 Nepārtrauktas datu uzlabošanas konvejera piemērs: datu apstrādes un analīzes plūsma būvniecības projektos.

Sistēma, kas apraksta vidēja lieluma uzņēmuma uzņēmējdarbības procesus, ir veidota pēc daudzīmeņu principa. Tā ietver: datu vākšanu, attīrišanu, apkopošanu, analītisko apstrādi un uz rezultātiem balstītu lēmumu pieņemšanu. Visus šos posmus mēs aplūkosim turpmāk grāmatā - gan teorētiskā kontekstā, gan ar praktisku piemēru pašdzību:

- Pirmajā līmenī **notiek datu ievadišana** (3.1-1. attēls). Informācija tiek saņemta gan manuāli (izmantojot pārskatus, veidlapas, žurnālus), gan automatizētā veidā (no API, sensoriem, programmatūras sistēmām). Dati var būt dažādas struktūras: ģeometriskie, teksta,

nestrukturēti. Šajā posmā ir nepieciešams standartizēt, strukturēt un unificēt informācijas plūsmas.

- Nākamais līmenis ir **datu apstrāde un pārveidošana**. Tas ietver datu attīrišanas, dublēšanās novēršanas, kļūdu labošanas un informācijas sagatavošanas turpmākai analīzei procesus (4.2-5. attēls). Šis posms ir ļoti svarīgs, jo analītikas kvalitāte ir tiesī atkarīga no datu tīrības un precizitātes.
- Pēc tam **dati tiek ievadīti specializētās tabulās, datu rāmjos vai datubāzēs**, kas sadalītas pa funkcionālajām jomām: budžeta plānošana un grāmatvedība, modeļi un rasējumi, logistika, drošība un infrastruktūra. Šāds sadalījums ļauj viegli piekļūt informācijai un veikt tās savstarpēju analīzi.
- Pēc tam dati tiek **apkopoti un attēloti analītiskajā paneļa panelī** (vitrīnā). Šeit tiek izmantotas aprakstošās, diagnostikas, prognozēšanas un preskriptīvās analīzes metodes. Tas palīdz atbildēt uz galvenajiem jautājumiem (1.1-4. attēls): kas notika, kāpēc tas notika, kas notiks nākotnē un kādas darbības ir jāveic. Piemēram, sistēma var noteikt kavējumus, prognozēt projektu pabeigšanu vai optimizēt resursus.
- Visbeidzot, pēdējā līmenī tiek veidoti **analītiski secinājumi un galvenie rādītāji**, kas palīdz uzraudzīt līgumu izpildi, pārvaldīt ieguldījumus un uzlabot uzņēmējdarbības procesus (7.4-2. attēls). Šī informācija kļūst par pamatu lēmumu pieņemšanai un uzņēmuma attīstības stratēģijai.

Līdzīgi arī dati no datu vākšanas tiek izmantoti stratēģiskajā vadībā. Turpmākajās grāmatas daļās mēs detalizēti aplūkosim katru posmu, pievēršot uzmanību datu veidiem, datu apstrādes metodēm, analīzes rīkiem un reāliem piemēriem, kā šīs pieejas tiek izmantotas būvniecības nozarē.

## Nākamie soļi: pārvērst haosu pārvaldāmā sistēmā

Šajā daļā mēs pētījām problēmas, kas saistītas ar informācijas "silosiem", un aplūkojām pārmērīgas sistēmas sarežģītības ietekmi uz uzņēmējdarbības rezultātiem, analizējot pāreju no ceturtās industriālās revolūcijas uz piekto, kurā galvenā nozīme ir datiem, nevis lietojumprogrammām. Mēs redzējām, kā izolētas informācijas sistēmas rada šķēršļus zināšanu apmaiņai, un pastāvīgā IT vides sarežģītība samazina produktivitāti un kavē inovācijas būvniecības nozarē.

Apkopojoš šo daļu, ir vērts uzsvērt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās pieejas ikdienas uzdevumos:

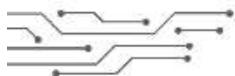
- Vizualizējet savu informācijas ainavu
  - Izveidojiet datu avotu vizuālo karti (Miro, Figma, Canva), ar kuriem regulāri strādājat.
  - Šajā kartē pievienojiet sistēmas un lietojumprogrammas, ko izmantojat savā darbā.
  - Identificēt potenciāli dublējošās funkcijas un liekos risinājumus.
  - Identificēt kritiskos punktus, kuros var rasties datu zudumi vai bojājumi datu pārraides laikā starp sistēmām.
- Personalizētu datu pārvaldības prakses īstenošana
  - Uzmanības pārvirzīšana no lietojumprogrammām uz datiem kā galveno procesu

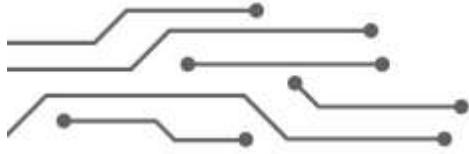
līdzekli.

- Datu avotu un apstrādes metodikas dokumentēšana, lai nodrošinātu pārredzamību
  - Izstrādāt mehānismus datu kvalitātes novērtēšanai un uzlabošanai
  - cestīes nodrošināt, lai dati tiktu ievadīti vienu reizi un izmantoti atkārtoti - tas ir efektīvas procesu organizācijas pamats.
- Veiciniet savā komandā uz datiem orientētu (uz datiem balstītu) pieeju.
- Ierosināt izmantot standartizētus un vienotus formātus savstarpējai datu apmaiņai.
  - Regulāri izvirziet jautājumus, kas saistīti ar datu kvalitāti un pieejamību, komandas sanāksmēs.
  - Iepazīstieties ar atvērtā koda alternatīvām rīkiem, ko izmantojat savu problēmu risināšanai.

Sāciet ar maziem elementiem - izvēlieties vienu konkrētu procesu vai datu kopumu, kas ir joti svarīgs jūsu darbā, un izmantojet uz datiem vērstu pieeju, pārorientējot uzmanību no rīkiem uz datiem. Sasniedzot panākumus vienā izmēģinājuma reizē, jūs ne tikai iegūsiet praktisku pieredzi, bet arī skaidri demonstrēsiet jaunās metodoloģijas priekšrocības savai komandai. Pildot lielāko daļu no šiem soļiem, ja jums rodas jautājumi, varat lūgt skaidrojumu un palīdzību jebkuram aktuālajam LLM.

Turpmākajās grāmatas daļās mēs detalizētāk aplūkosim datu strukturēšanas un saskaņošanas metodes un izpētīsim praktiskas piejas neviendabīgas informācijas integrēšanai. Īpaša uzmanība tiks pievērsta pārejai no atšķirīgiem datu silosiem uz vienotām datu ekosistēmām, kam ir būtiska nozīme būvniecības nozares digitālās transformācijas procesā.





### III DAĻA

## DATU SISTĒMA BŪVΝIECĪBAS UZNĒMĒJDARBĪBAS PROCESOS

Trešajā daļā tiek veidota visaptveroša izpratne par datu tipoloģiju būvniecībā un to efektīvas organizācijas metodēm. Tieks analizētas strukturētu, nestrukturētu, daļēji strukturētu, teksta un ģeometrisko datu īpatnības un darba specifika būvniecības projektu kontekstā. Apskatīti mūsdienu datu uzglabāšanas formāti un protokoli informācijas apmaiņai starp dažādām nozarē izmantotajām sistēmām. Aprakstīti praktiski rīki un paņemieni daudzformātu datu pārveidošanai vienotā strukturētā vidē, tostarp, kā integrēt CAD (BIM) datus. Tieks piedāvātas piejas, kā nodrošināt datu kvalitāti, izmantojot standartizāciju un validāciju, kas ir ļoti svarīga būvniecības aprēķinu precizitātei. Detalizēti analizēti mūsdienu tehnoloģiju (Python Pandas, LLM modeļi) izmantošanas praktiskie aspekti ar koda piemēriem, lai risinātu tipiskas būvniecības nozares problēmas. Ir pamatota kompetences centra (CoE) kā organizatoriskas struktūras nozīme informācijas pārvaldības pieju koordinēšanai un standartizācijai.

## NODAĻA 3.1.

# DATU TIPI BŪVNIECĪBĀ

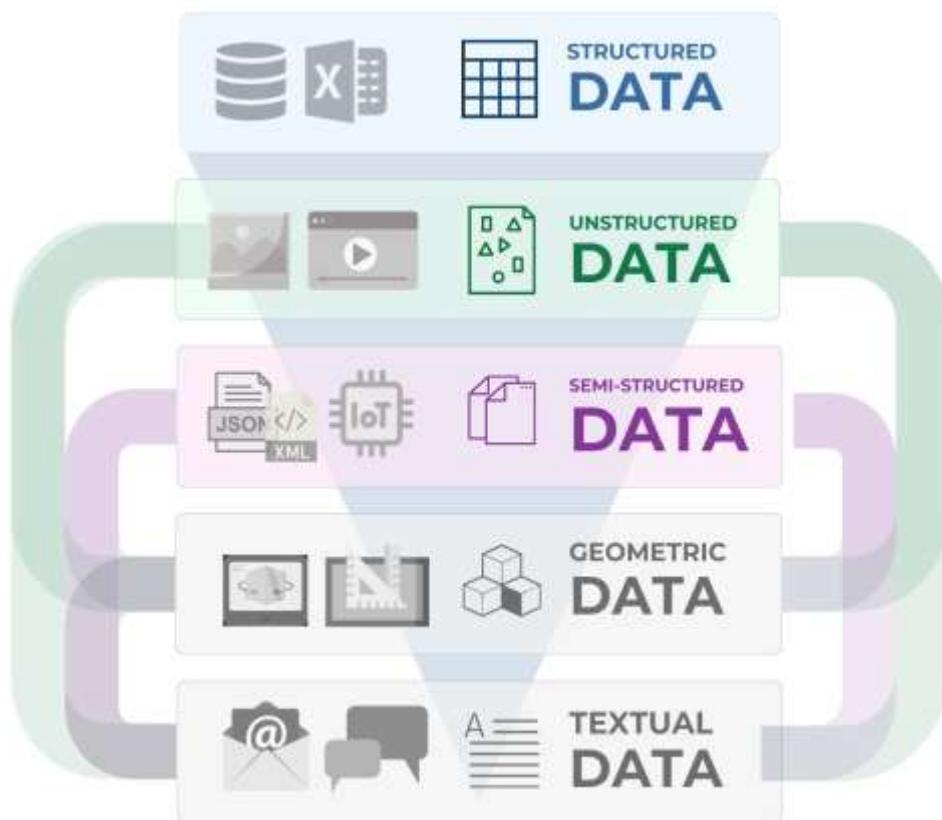
### Svarīgākie datu veidi būvniecības nozarē

Mūsdieni būvniecības nozarē uzņēmumu sistēmas, lietojumprogrammas un datu noliktavas tiek aktīvi pildītas ar dažāda veida un formāta informāciju un datiem ( 3.1-1. attēls). Aplūkosim sīkāk galvenos datu veidus, kas veido mūsdieni būvniecības nozarē strādājoša uzņēmuma informācijas ainavu:

- **Strukturēti** dati: šiem datiem ir skaidra organizatoriskā struktūra, piemēram, Excel izklājlapas un relāciju datu bāzes.
- **Nestrukturizēti** dati: tā ir informācija, kas nav sakārtota saskaņā ar stingriem noteikumiem. Šādu datu piemēri ir teksts, video, fotogrāfijas un audioieraksti.
- **Vāji strukturēti** dati: šie dati ir starpposms starp strukturētiem un nestrukturētiem datiem. Tajos ir struktūras elementi, bet šī struktūra ne vienmēr ir skaidra vai bieži vien ir aprakstīta dažādās shēmās. Dalēji strukturētu datu piemēri būvniecībā ir tehniskās specifikācijas, projekta dokumentācija vai progresu ziņojumi.
- **Teksta dati**: ietver visu, kas iegūts no mutiskas un rakstiskas saziņas, piemēram, e-pasta vēstules, sanāksmju un tikšanos stenogrammas.
- **Ģeometriskie** dati: šie dati nāk no CAD programmām, kurās speciālisti izveido projekta elementu ģeometriskos datus vizualizācijai, tilpuma vērtību apstiprināšanai vai sadursmju pārbaudei.

Iz svarīgi atzīmēt, ka ģeometriskie un teksta (burtciparu) dati nav atsevišķa kategorija, bet var būt sastopami visos trīs datu veidos. Piemēram, ģeometriskie dati var būt daļa gan no strukturētiem datiem (parametriskie CAD formāti), gan nestrukturētiem datiem (skenēti rasējumi). Līdzīgi arī teksta dati var būt gan organizēti datubāzēs (strukturēti dati), gan pastāvēt kā dokumenti bez skaidras struktūras.

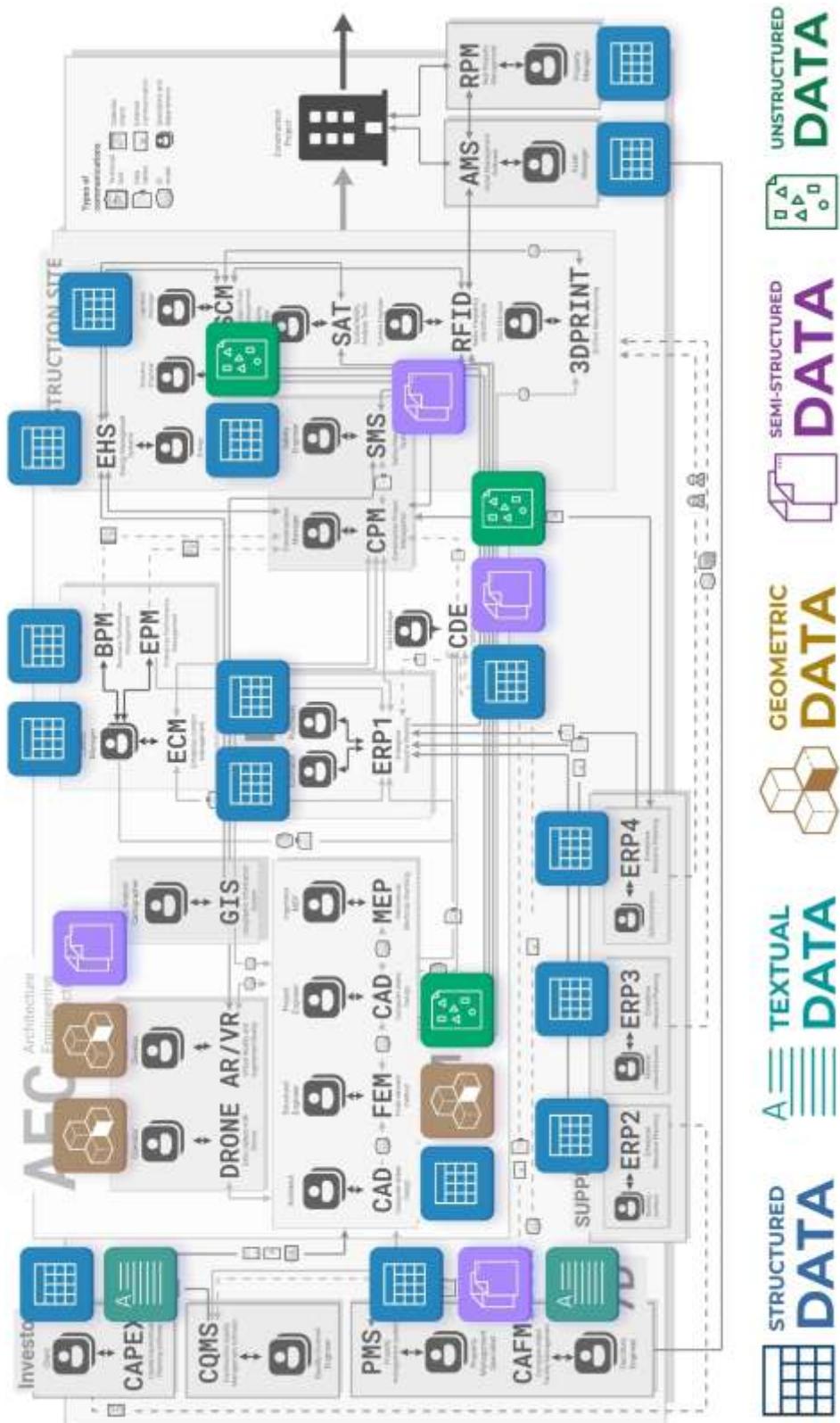
Katrs datu veids būvniecības uzņēmumā ir unikāls elements uzņēmuma informācijas aktīvu mozaīkā. No nestrukturētiem datiem, piemēram, būvlaukumu attēliem un sanāksmju audioierakstiem, līdz strukturētiem ierakstiem, tostarp tabulām un datubāzēm, katram elementam ir svarīga loma uzņēmuma informācijas ainavas veidošanā.



Attēls 3.1-1 Inženieriem un datu pārvaldītājiem ir jāiemācās strādāt ar visiem būvniecības nozarē izmantotajiem datu veidiem.

Šeit ir sniegs tikai dažu būvniecībā izmantoto sistēmu un saistīto datu tipu paraugs ( 3.1-2. attēls):

- **ERP** (Enterprise Resource Planning) - apstrādā vispārēji strukturētus datus, lai palīdzētu pārvaldīt uzņēmuma resursus un integrēt dažādus uzņēmējdarbības procesus.
- **CAD** (Computer-Aided Design) apvienojumā ar **BIM** (Building Information Modeling) - izmanto ģeometriskus un daļēji strukturētus datus, lai projektētu un modelētu būvniecības projektus, nodrošinot informācijas precizitāti un konsekvenči projektēšanas posmā.
- **GIS** (ģeogrāfiskās informācijas sistēmas) - strādā ar ģeometriskiem un strukturētiem datiem, lai veidotu un analizētu kartogrāfiskus datus un telpiskās attiecības.
- **RFID** (radiofrekvenču identifikācija) - izmanto daļēji strukturētus datus, lai efektīvi izsekotu materiālus un iekārtas būvlaukumā, izmantojot radiofrekvenču identifikāciju.
- **ECM** (inženiertehniskā satura pārvaldība) ir sistēma inženiertehnisko datu un dokumentācijas, tostarp daļēji strukturētu un nestukturētu datu, piemēram, tehnisko rasējumu un projektēšanas dokumentu, pārvaldībai.

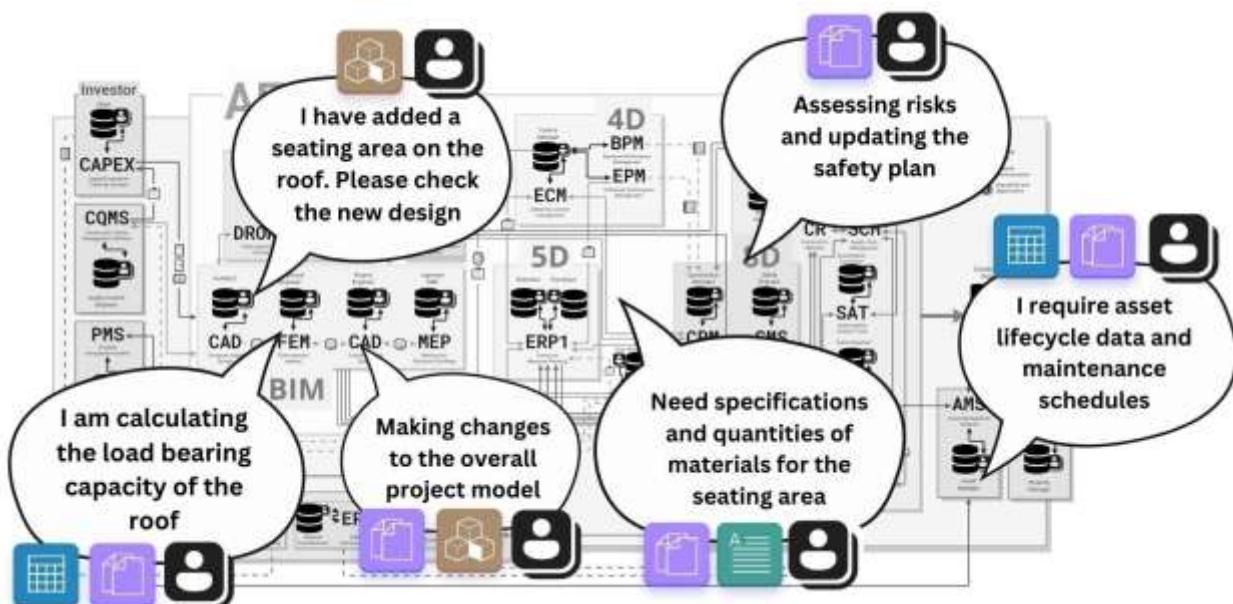


Attēls 3.1-2 Dažādās sistēmās tiek izmantoti dažādi formāti un dati, kas ir jātransponē sarežģītai integrācijai piemērotā formā.

Šīs un daudzas citas uzņēmuma sistēmas pārvalda plašu datu klāstu, sākot no strukturētiem tabulas datiem līdz sarežģītiem geometriskiem modeļiem, nodrošinot integrētu sadarbību projektēšanas, plānošanas un būvniecības vadības procesos.

Vienkāršota dialoga piemērā (3.1-3. attēls) redzams, ka būvniecības projekta speciālisti apmainās ar dažāda veida datiem:

- ⌚ **Arhitekts:** "Nemot vērā klienta vēlmes, es uz jumta ierīkoju atpūtas zonu. Lūdzu, aplūkojiet jauno projektu" (geometriskie dati - modelis).
- ⌚ **Būvinženieris:** "Projekts ir saņemts. Es aprēķinu jumta nestspēju jaunajai atpūtas zonai" (strukturēti un daļēji strukturēti dati - aprēķinu tabulas).
- ⌚ **Iepirkumu vadītājs:** "Nepieciešamas atpūtas zonas materiālu specifikācijas un daudzumi, lai organizētu iepirkumu" (teksta un daļēji strukturēti dati - saraksti un specifikācijas).
- ⌚ **Veselības un drošības inženieris:** "Saņemti dati par jauno zonu. Izvērtēju riskus un atjauninu drošības plānu" (daļēji strukturēti dati - dokumenti un plāni).
- ⌚ **Speciālists BIM - modelēšana:** "Izmainu veikšana kopējā projekta modelī, lai pielāgotu darba dokumentāciju" (geometriskie dati un daļēji strukturēti dati).
- ⌚ **Projekta vadītājs:** "Es iekļauju jauno atpūtas zonu darba grafikā. Es atjauninu grafikus un resursus projekta vadības sistēmā" (strukturēti un daļēji strukturēti dati - grafiki un plāni).
- ⌚ **Iekārtu uzturēšanas speciālists (FM):** "Es sagatavoju datus turpmākai atpūtas zonas uzturēšanai un ievadu tos īpašumu pārvaldības sistēmā" (strukturēti un daļēji strukturēti dati - instrukcijas un uzturēšanas plāni).



Attēls 3.1-3 Komunikācija starp speciālistiem notiek gan teksta, gan datu līmenī.

Ktrs speciālists strādā ar dažāda veida datiem, lai nodrošinātu efektīvu komandas sadarbību un projekta panākumus. Izpratne par atšķirībām starp strukturētiem, daļēji strukturētiem un nestukturētiem datiem ļauj apzināties, kāda unikāla loma katram datu veidam ir digitālajos biznesa procesos. Ir svarīgi ne tikai zināt, ka pastāv dažādi datu veidi, bet arī saprast, kā, kur un kāpēc tie tiek izmantoti.

Vēl pavisam nesen ideja apvienot tik dažādus datus šķita ambicioza, bet grūti īstenojama. Šodien tā jau ir daļa no ikdienas prakses. Dažādu shēmu un struktūru datu integrācija ir kļuvusi par mūsdienu informācijas sistēmu arhitektūras neatņemamu sastāvdalī.

Turpmākajās nodalās mēs detalizēti aplūkosim galvenos standartus un piejas, kas ļauj strukturētus, daļēji strukturētus un nestukturētus datus apvienot vienotā saskaņotā pārskatā. Īpaša uzmanība tiks pievērsta strukturētiem datiem un relāciju datubāzēm kā galvenajiem informācijas glabāšanas, apstrādes un analīzes mehānismiem būvniecības nozarē.

## Strukturēti dati

Būvniecības nozarē informācija nāk no daudziem avotiem - rasējumiem, specifikācijām, grafikiem un ziņojumiem. Lai efektīvi pārvaldītu šo informācijas plūsmu, tā ir jāstrukturē. Strukturēti dati ļauj jums sakārtot informāciju ērtā, viegli lasāmā un pieejamā formā.

Saskaņā ar JB Knowledge 5. ikgadējo būvniecības tehnoloģiju ziņojumu [17] 67% būvniecības projektu vadības profesionāļu seko līdzi un novērtē darba izpildi manuāli vai izmantojot izklājlapas.

Daži no izplatītākajiem strukturēto datu formātiem ir XLSX un CSV. Tos plaši izmanto informācijas glabāšanai, apstrādei un analīzei izklājlapās. Šādās izklājlapās dati ir attēloti kā rindas un kolonas, tāpēc tos ir viegli lasīt, rediģēt un analizēt.

**XLSX**, kas ir Microsoft izveidots **formāts**, kura pamatā ir XML -struktūras izmantošana, un tas tiek arhivēts, izmantojot ZIP algoritmu. Formāta galvenās iezīmes:

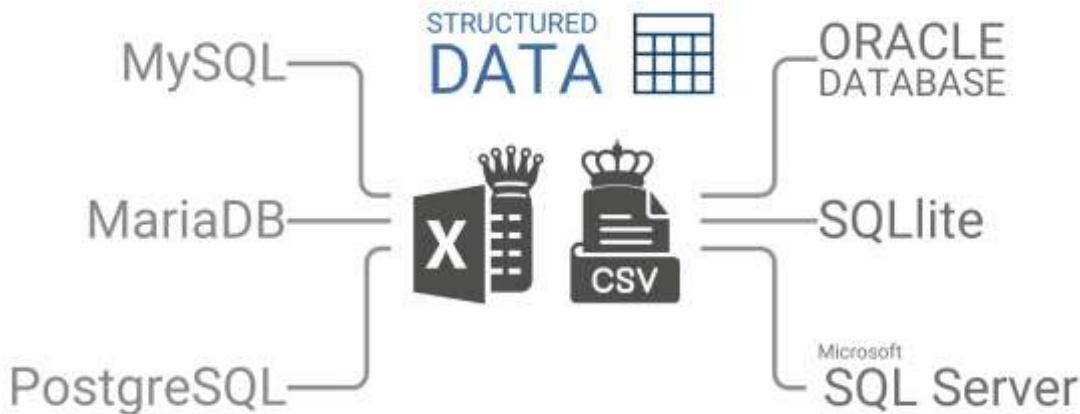
- Sarežģītu formulu, diagrammu un makru atbalsts.
- Iespēja saglabāt datus dažādās lapās, kā arī formatēt informāciju.
- Optimizēts Microsoft Excel, bet ir saderīgs arī ar citiem biroja komplektiem.

**CSV formāts** ir vienkārša teksta fails, kurā vērtības ir atdalītas ar komatiem, semikoloniem vai citām norobežojošām zīmēm. Galvenās priekšrocības:

- Universāla savietojamība ar dažādām programmām un operētājsistēmām.
- Viegli importēt/eksportēt uz datu bāzēm un analītiskajām sistēmām.
- Vieglā apstrāde pat teksta redaktoriem.

Tomēr CSV neatbalsta formulas un formatēšanu, tāpēc tā galvenais pielietojums ir datu apmaiņa starp sistēmām un masveida informācijas atjaunināšana. Tā kā CSV ir universāls un platformas neatkarīgs, tas ir kļuvis par populāru rīku datu pārsūtīšanai heterogēnās IT vidēs.

Divi formāti XLSX un CSV darbojas kā saikne starp dažādām sistēmām, kas strādā ar strukturētiem datiem ( 3.1-4. attēls). Tie ir īpaši noderīgi uzdevumos, kur svarīga ir lasāmība, manuāla rediģēšana un pamata savietojamība



**Attēls 3.1-4 XLSX un CSV formāti ir saikne starp dažādām sistēmām, kas strādā ar strukturētiem datiem.**

Platformas neatkarība padara CSV par populārāko formātu datu pārsūtīšanai heterogēnās IT vidēs un sistēmās.

Tomēr XLSX un CSV nav paredzēti augstas veikspējas skaitļošanai vai liela datu apjoma ilgtermiņa glabāšanai. Šādiem nolūkiem izmanto modernākus struktureņus formātus, piemēram, Apache Parquet, Apache ORC, Feather, HDF5. Šie formāti sīkāk tiks aplūkoti šīs grāmatas 9. daļas nodaļā "Lielo datu glabāšana: populāru formātu analīze un to efektivitāte"

Praksē Excel ar XLSX formātu biežāk tiek izmantots nelielu uzdevumu veikšanai un rutīnas procesu automatizēšanai. Sarežģītākiem scenārijiem ir jāizmanto datu pārvaldības sistēmas, piemēram, ERP, PMIS CAFM, CPM, SCM un citas ( 3.2-1. attēls). Šajās sistēmās tiek glabāti strukturēti dati, uz kuriem balstās uzņēmuma informācijas plūsmu organizācija un pārvaldība.

Mūsdienu datu pārvaldības informācijas sistēmas, ko izmanto būvniecības nozarē, balstās uz strukturētiem datiem, kas sakārtoti tabulu veidā. Lai nodrošinātu uzticamu, mērogojamu un holistisku liela apjoma informācijas pārvaldību, lietojumprogrammu un sistēmu izstrādātāji izmanto relāciju datubāzu pārvaldības sistēmas (RDBMS).

## Relāciju datu bāzes RDBMS un SQL vaicājumu valoda

Lai efektīvi uzglabātu, apstrādātu un analizētu datus, **relacionālās datubāzes (RDBMS)** ir datu glabāšanas sistēmas, kas informāciju sakārto tabulās ar noteiktām attiecībām starp tām.

Datubāzēs (RDBMS) organizēti dati nav tikai digitāla informācija; tie ir pamats darījumiem un mijiedarbībai starp dažādām sistēmām.

Tālāk ir uzskaitītas dažas visbiežāk sastopamās relāciju datubāzu pārvaldības sistēmas (RDBMS) (3.1-5. attēls):

- **MySQL** (Open Source) ir viena no populārākajām RDBMS, kas ir daļa no LAMP (Linux, Apache, MySQL, PHP /Perl/Python). To plaši izmanto tīmekļa vietņu izstrādē, jo tā ir vienkārša un augstas veiktspējas.
- **PostgreSQL** (Open Source) ir jaudīga objekt-relatīvā sistēma, kas pazīstama ar savu uzticamību un uzlabotām funkcijām. Tā ir piemērota sarežģītiem uzņēmumu risinājumiem.
- **Microsoft SQL Server** ir Microsoft komerciāla sistēma, ko plaši izmanto korporatīvajās vidēs, pateicoties tās integrācijai ar citiem uzņēmuma produktiem un augstajam drošības līmenim.
- **Oracle datubāze** ir viena no jaudīgākajām un uzticamākajām DBVS, ko izmanto lielos uzņēmumos un kritiski svarīgās lietojumprogrammās.
- **IBM DB2** - paredzēta lielām korporācijām, nodrošina augstu veiktspēju un klūdu toleranci.
- **SQLite** (Open Source) ir viegla iegultā datubāze, kas ideāli piemērota mobilajām lietojumprogrammām un autonomām sistēmām, piemēram, CAD projektēšanas programmatūrai (BIM).

Būvniecības nozarē populārākās datubāzu pārvaldības sistēmas - MySQL, PostgreSQL, Microsoft SQL Server, Oracle® Database, IBM® DB2 un SQLite - strādā ar strukturētiem datiem. Visas šīs DBVS ir jaudīgi un elastīgi risinājumi dažādu uzņēmējdarbības procesu un lietojumprogrammu pārvaldībai, sākot no nelielām tīmekļa vietnēm un beidzot ar liela mēroga uzņēmumu sistēmām (3.2-1. attēls).

Saskaņā ar Statista [48] datiem 2022. gadā relāciju datubāzu pārvaldības sistēmas (RDBMS) veidos aptuveni 72% no visām izmantotajām DBVS.

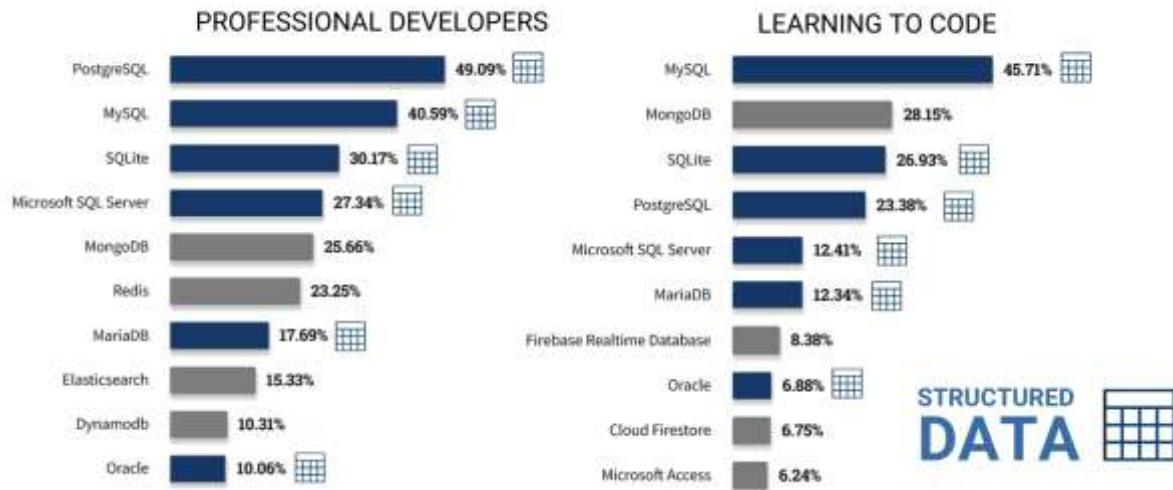
	Rank			DBMS	Database Model	Open Source vs Commercial
	Mar2025	Feb2025	Mar2024			
	1.	1.	1.	Oracle®	Relational, Multi-model	Commercial
	2.	2.	2.	MySQL	Relational, Multi-model	Open Source
	3.	3.	3.	Microsoft® SQL Server	Relational, Multi-model	Commercial
	4.	4.	4.	PostgreSQL	Relational, Multi-model	Open Source
	5.	5.	5.	MongoDB	Document, Multi-model	Open Source
	6.	7.	9.	Snowflake®	Relational	Commercial
	7.	6.	6.	Redis®	Key-value, Multi-model	Open Source
	8.	8.	7.	Elasticsearch®	Multi-model	Open Source
	9.	9.	8.	IBM Db2	Relational, Multi-model	Commercial
	10.	10.	10.	SQLite	Relational	Open Source
	11.	11.	12.	Apache Cassandra®	Multi-model	Open Source
	12.	12.	11.	Microsoft Access®	Relational	Open Source
	13.	13.	17.	Databricks®	Multi-model	Commercial
	14.	14.	13.	MariaDB	Relational, Multi-model	Open Source
	15.	15.	14.	Splunk	Search engine	Commercial
	16.	16.	16.	Amazon DynamoDB	Multi-model	Commercial
	17.	17.	15.	Microsoft Azure SQL	Relational, Multi-model	Commercial

Attēls 3.1-5 Strukturētu datubāzu (atzīmētas zilā krāsā) izmantošanas popularitāte DBVS ranžēšanā (pēc [49]).

Atvērtā pirmkoda datubāzes ir diezgan viegli instalēt - pat bez plašām tehniskām zināšanām. Atvērtā pirmkoda sistēmas, piemēram, PostgreSQL, MySQL vai SQLite, ir pieejamas bez maksas un darbojas lielākajā daļā operētājsistēmu: Windows, macOS un Linux. Viss, kas jums nepieciešams, ir doties uz projekta oficiālo tīmekļa vietni, lejupielādēt instalēšanas programmu un sekot norādījumiem. Vairumā gadījumu instalēšana aizņem ne vairāk kā 10-15 minūtes. Mēs modelēsim un izveidosim vienu šādu datubāzi grāmatas ceturtajā daļā ( 4.3-8. attēls).

Ja jūsu uzņēmums izmanto mākoņpakaļpojumus (piemēram, Amazon Web Services, Google Cloud vai Microsoft Azure), datubāzi varat izvietot ar pāris klikšķiem - platforma piedāvās instalēšanai gatavus veidnes. Pateicoties koda atvērtībai, šādas datubāzes ir viegli pielāgot saviem uzdevumiem, un milzīgā lietotāju kopiena vienmēr palīdzēs jums atrast risinājumu jebkurai problēmai.

RDBMS joprojām ir pamats daudzām biznesa lietojumprogrammām un analīzes platformām (3.1.1-6. attēls), kas ļauj uzņēmumiem efektīvi uzglabāt, apstrādāt un analizēt datus un tādējādi pieņemt pamatotus un savlaciņgus lēmumus.



Attēls 3.1-6 Lielākā IT foruma StackOverFlow izstrādātāju aptauja par to, kuras datubāzes viņi izmantoja pagājušajā gadā un kuras vēlas izmantot nākamajā gadā (RDBMS ir izceltas zilā krāsā) (pamatojoties uz [50]).

RDBMS nodrošina uzticamību, datu konsekvenci, transakciju atbalstu un izmanto jaudīgu vaicājumu valodu - SQL (strukturēto vaicājumu valodu), kas bieži tiek izmantota analīzē un ļauj viegli iegūt, modifīcēt un analizēt datubāzēs glabāto informāciju. SQL ir galvenais rīks darbam ar datiem relāciju sistēmās.

## SQL - vaicājumi datubāzēs un jaunas tendences

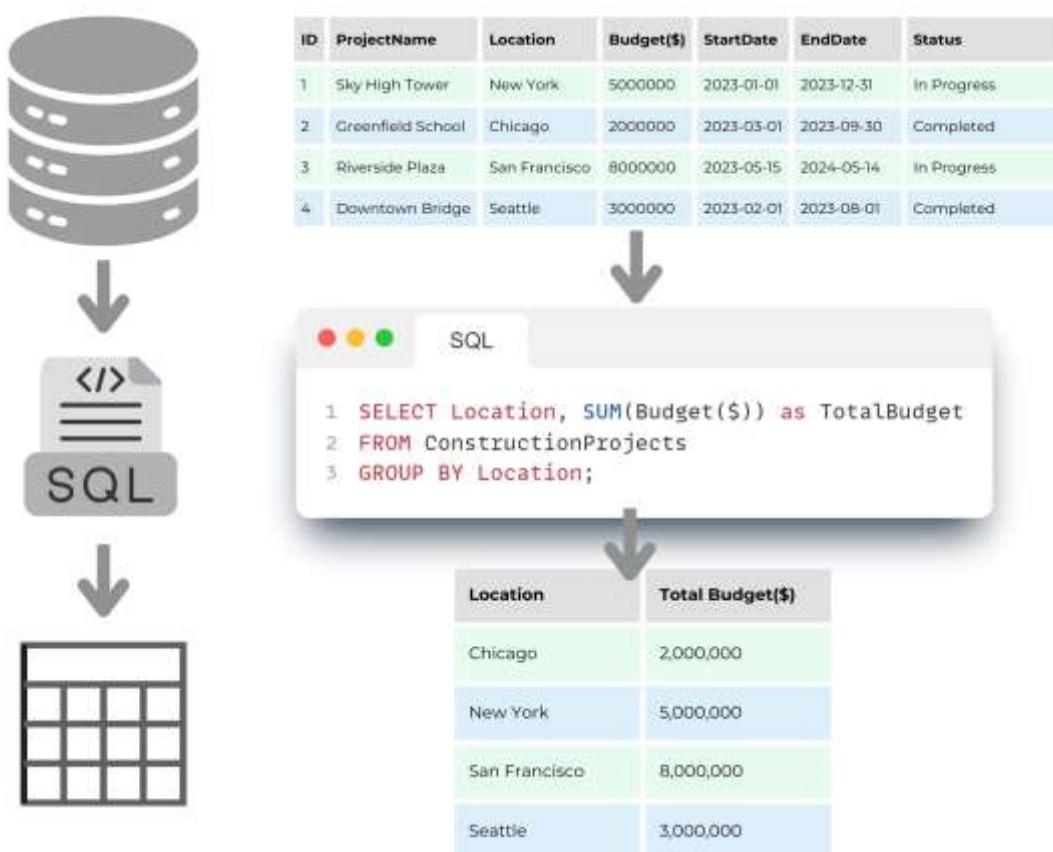
SQL valodas, ko bieži izmanto relāciju datubāzēs, galvenā priekšrocība salīdzinājumā ar citiem informācijas pārvaldības veidiem (piemēram, izmantojot klasiskās Excel izklājlapas) ir ļoti lielu datu bāzu apjomu atbalsts ar lielu pieprasījumu apstrādes ātrumu.

Strukturētā vaicājumu valoda (SQL) ir specializēta programmēšanas valoda, kas paredzēta informācijas glabāšanai, apstrādei un analīzei relāciju datubāzēs. SQL izmanto datu izveidei, pārvaldībai un piekļuvei tiem, ļaujot efektīvi atrast, filtrēt, apvienot un apkopot informāciju. Tā kalpo kā galvenais rīks piekļuvei datiem, nodrošinot ērtu un formalizētu veidu, kā mijiedarboties ar informācijas krātuvēm.

SEQUEL-SQL sistēmu evolūcija iet cauri tādiem nozīmīgiem produktiem un uzņēmumiem kā Oracle, IBM DB2, Microsoft SQL Server, SAP, PostgreSQL un MySQL, un kulminē SQLite un MariaDB [51]. SQL nodrošina izklājlapas iespējas, kas nav atrodamas Excel, padarot datu apstrādi mērogojamāku, drošāku un viegli automatizējamu:

- **Datu struktūru izveide un pārvaldība (DDL):** SQL var izveidot, modifīcēt un dzēst datubāzes tabulas, izveidot saites starp tām un definēt datu glabāšanas struktūras. Savukārt Excel strādā ar fiksētām lapām un šūnām, bez skaidri definētām attiecībām starp lapām un datu kopām.

- **Datu manipulācijas (DML):** SQL ļauj masveidā pievienot, modifcēt, dzēst un iegūt datus lielā ātrumā, veicot sarežģītus pieprasījumus ar filtrēšanu, šķirošanu un tabulu apvienošanu (3.1-7. attēls). Liela apjoma informācijas apstrādei programmā Excel ir jāveic manuālas darbības vai jāizmanto īpaši makroekrāni, kas palēnina procesu un palielina kļūdu iespējamību.
- **Piekļuves kontrole (DCL):** SQL ļauj diferencēt dažādu lietotāju piekļuves tiesības datiem, ierobežojot iespējas rediģēt vai skatīt informāciju. Savukārt programmā Excel piekļuve ir vai nu koplietojama (pārsūtot failu), vai arī nepieciešami sarežģīti iestatījumi ar atļauju koplietošanu, izmantojot mākoņpakalpojumus.



Attēls 3.1-7 DML piemērs SQL: ātra apstrāde, grupēšana un apkopošana ar dažām koda rindiņām automātiskai datu apstrādei.

Excel atvieglo darbu ar datiem, pateicoties vizuālai un intuitīvai struktūrai. Tomēr, palielinoties datu apjomam, Excel veiktspēja samazinās. Excel ir arī ierobežots datu apjoms, ko tas var uzglabāt, - maksimums ir viens miljons rindu, un veiktspēja pasliktinās ilgi pirms šī ierobežojuma sasniegšanas. Tāpēc, gan Excel šķiet piemērotāks neliela datu apjoma vizualizēšanai un apstrādei, SQL ir labāk piemērots lielu datu kopu apstrādei.

Nākamais strukturēto datu attīstības posms bija kolonnveida datubāzu (Columnar Databases) rašanās, kas ir alternatīva tradicionālajām relāciju datubāzēm, īpaši, ja runa ir par ievērojami lielākiem datu apjomiem un analītiskiem aprēķiniem. Atšķirībā no rindu datubāzēm, kurās dati tiek saglabāti rindiņa pēc rindas, kolonnveida datubāzēs informācija tiek reģistrēta pa kolonnām. Salīdzinājumā ar

Klasiskajām datubāzēm tas ļauj:

- Samaziniet glabāšanas vietu, efektīvi saspiežot viendabīgus datus kolonnās.
- Paātriniet analītiskos vaicājumus, jo tiek nolasīti tikai nepieciešamie kolonas, nevis visa tabula.
- Lielo datu optimizācija un datu noliktavas, piemēram, Data Lakehouse Architecture.

Par kolonnesveida datu bāzēm, Pandas DataFrame, Apache Parquet, HDF5, kā arī par uz tiem balstītu lielo datu -glabātavu izveidi datu analīzes un apstrādes vajadzībām plašāk runāsim turpmākajās šīs grāmatas nodaļās - "DataFrame: universāls tabulāro datu formāts" un "Datu glabāšanas formāti un darbs ar Apache Parquet: DWH -datu noliktavas un datu ezeru arhitektūra".

## Nestrukturizēti dati

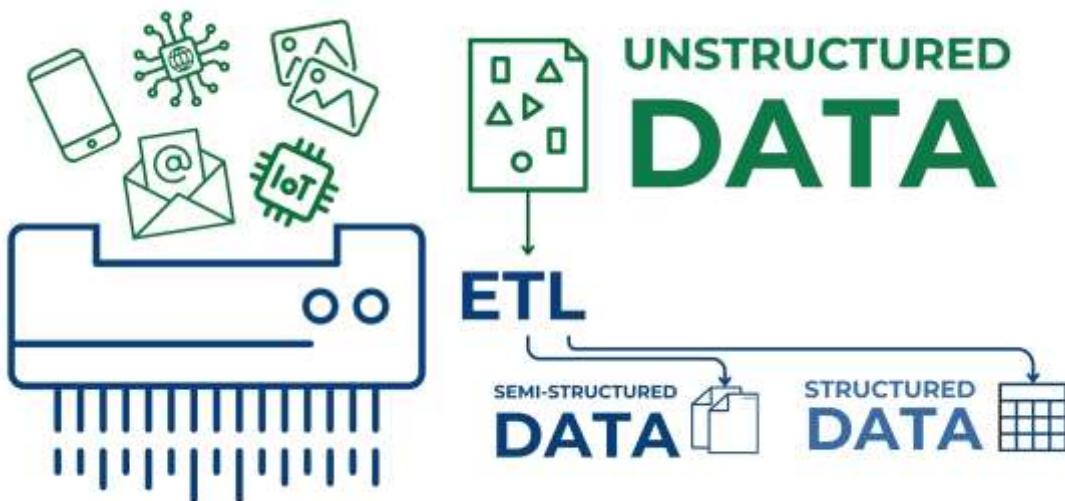
Lai gan lielākā daļa lietojumprogrammās un informācijas sistēmās izmantoto datu ir strukturētā formā, lielākā daļa būvniecībā radītās informācijas ir nestrukturēti dati - attēli, video, teksta dokumenti, audioieraksti un cita veida saturs. Tas jo īpaši attiecas uz būvniecības, ekspluatācijas un tehniskās uzraudzības posmiem, kur dominē vizuālā un teksta informācija.

**Nestrukturizēti dati** ir informācija, kurai nav iepriekš definēta modeļa vai struktūras, kas nav sakārtota tradicionālajās rindās un kolonnās kā datu bāzēs vai tabulās.

Kopumā nestrukturētus datus var iedalīt divās kategorijās:

- Cilvēka radīti nestrukturēti dati, kas ietver dažādus cilvēka radīta satura veidus: teksta dokumentus, e-pasta vēstules, attēlus, videoklipus utt.
- Mašīnu ģenerētus nestrukturētus datus rada ierīces un sensori: tie ietver, piemēram, žurnālu datnes, GPS datus, lietu interneta rezultātus (IoT) un citu telemetrijas informāciju no būvlaukuma.

Atšķirībā no strukturētiem datiem, kurus ir ērti sakārtot tabulās un datubāzēs, nestrukturētiem datiem pirms to integrēšanas informācijas sistēmās ir nepieciešami papildu apstrādes posmi ( 3.1-8. attēls). Tehnoloģiju izmantošana šādu datu automatizētai vākšanai, analīzei un pārveidošanai paver jaunas iespējas uzlabot būvniecības efektivitāti, samazināt klūdas un līdz minimumam samazināt cilvēciskā faktora ietekmi.



Attēls 3.1-8 Nestukturētu datu apstrāde sākas ar to pārveidošanu daļēji strukturētos un strukturētos datos.

Nestukturizēti dati veido līdz pat 80% no visas informācijas [52], ar ko saskaras profesionāļi uzņēmumos, tāpēc turpmākajās grāmatas nodaļās mēs detalizēti un ar piemēriem aplūkosim to veidus un apstrādi.

Diskusijas ērtības labad teksta dati ir iedalīti atsevišķās kategorijās. Lai gan tie ir diezgan nestukturētu datu veids, to nozīme un izplatība būvniecības nozarē prasa īpašu uzmanību.

### Teksta dati: starp nestukturētu haosu un strukturētu haosu y

Teksta dati būvniecības nozarē aptver plašu informācijas formātu un veidu klāstu, sākot ar papīra dokumentiem un beidzot ar neformālām saziņas metodēm, piemēram, vēstulēm, sarunām, darba saraksti un mutiskām sanāksmēm būvlaukumā. Visi šie teksta dati ietver svarīgu informāciju būvniecības projektu vadībai, sākot ar informāciju par projektēšanas lēmumiem un izmaiņām plānos un beidzot ar diskusijām par drošības jautājumiem un sarunām ar darbuzņēmējiem un klientiem (3.1-9. attēls).



Attēls 3.1-9 Teksta dati, kas ir viens no populārākajiem informācijas veidiem, ko izmanto saziņā starp projekta dalībniekiem.

Teksta informācija var būt gan formalizēta, gan nestukturēta. Formalizēti dati ietver Word dokumentus (.doc,.docx), PDF, kā arī sanāksmju protokolu teksta failus (.txt). Neformalizēti dati ietver ziņojumapmaiņas un e-pasta saraksti, sanāksmju stenogrammas (Teams, Zoom, Google Meet) un diskusiju audioierakstus (.mp3,.wav), kurus nepieciešams konvertēt tekstā.

Taču, lai gan rakstiskiem dokumentiem, piemēram, oficiāliem pieprasījumiem, līguma noteikumiem un nosacījumiem un e-pasta vēstulēm, parasti jau ir zināma struktūra, mutiskā saziņa un darba sarakste bieži vien ir nestukturēta, tāpēc tos ir grūti analizēt un integrēt projektu vadības sistēmās.

Teksta datu efektīvas pārvaldības atslēga ir pārvērst tos strukturētā formātā. Tas jauj apstrādāto informāciju automātiski integrēt esošajās sistēmās, kas jau strādā ar strukturētiem datiem.



Attēls 3.1-10 Teksta satura pārvēršana strukturētos datos.

Lai efektīvi izmantotu teksta informāciju, tā automātiski jāpārveido strukturētā formā ( 3.1-10. attēls). Šis process parasti ietver vairākus posmus:

- **Teksta atpazīšana (OCR)** - dokumentu un rasējumu attēlu pārvēršana mašīnlasāmā formātā.
- **Teksta analīze (NLP)** - automātiska galveno parametru (datumi, summas un skaitļi, kas attiecas uz projektu) noteikšana.
- **Datu klasifikācija** - informācijas kategorizēšana (finances, logistika, riska vadība).

Pēc atpazīšanas un klasifikācijas jau strukturētus datus var integrēt datubāzēs un izmantot automatizētās ziņošanas un pārvaldības sistēmās.

## Dalēji strukturēti un brīvi strukturēti dati

Dalēji strukturēti dati ietver zināmu organizācijas līmeni, bet tiem nav stingras shēmas vai struktūras. Lai gan šāda informācija ietver strukturētus elementus (piemēram, datumus, darbinieku vārdus un izpildīto uzdevumu sarakstus), tās noformējums var ievērojami atšķirties atkarībā no projekta vai pat atkarībā no darbinieka. Šādu datu piemēri ir darba laika uzskaites žurnāli, progresu ziņojumi un grafiki, kas var tikt sniegti dažādos formātos.

Dalēji strukturētus datus ir vieglāk analizēt nekā nestukturētus datus, taču, lai tos integrētu standartizētās projektu vadības sistēmās, nepieciešama papildu apstrāde.

Darbs ar daļēji strukturētiem datiem, kuriem raksturīga pastāvīgi mainīga struktūra, rada ievērojamus izaicinājumus. Tas ir tāpēc, ka datu struktūras mainīgums prasa atsevišķu individuālu pieeju katra daļēji strukturētu datu avota apstrādei un analīzei.

Taču, kamēr darbs ar nestukturētiem datiem prasa daudz pūļu, daļēji strukturētu datu apstrādi var veikt ar salīdzinoši vienkāršām metodēm un rīkiem.

**Vāji strukturēti** dati ir vispārīgāks termins, kas apzīmē datus ar minimālu vai nepilnīgu struktūru. Visbiežāk tie ir teksta dokumenti, tērzēšanas, e-pasta vēstules, kur atrodami daži metadati (piemēram, datums, sūtītājs), bet lielākā daļa informācijas ir sniepta haotiski.

Būvniecībā brīvi strukturēti dati ir sastopami dažādos procesos. Piemēram, tie var ietvert:

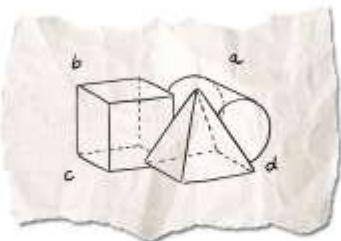
- Aplēses un cenu piedāvājumi - tabulas ar datiem par materiāliem, daudzumiem un izmaksām, bet bez vienota formāta.
- Rasējumi un inženierhēmas - faili PDF formātā vai DWG formātā, kuros ir teksta anotācijas un metadati, bet nav stingri noteiktas struktūras.
- Darba grafiki - dati no MS Project, Primavera P6 vai citām sistēmām, kurām var būt atšķirīga eksporta struktūra.
- CAD (BIM -modelji) - satur konstrukcijas elementus, bet datu attēlojums ir atkarīgs no programmatūras un projekta standarta.

Ģeometriskos datus, ko sagatavo CAD sistēmas, var klasificēt tāpat kā daļēji strukturētus datus. Tomēr mēs klasificēsim ģeometriskos CAD (BIM) datus kā atsevišķu datu veidu, jo tos, tāpat kā teksta datus, bieži vien uzņēmuma procesos var uzskatīt par atsevišķu datu veidu.

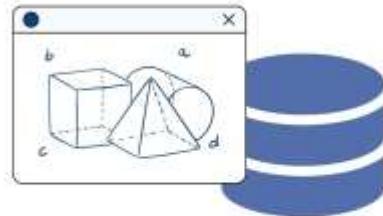
## Ģeometriskie dati un to izmantošana

Ja metadati par projekta elementiem gandrīz vienmēr tiek glabāti tabulu, strukturētu vai brīvi strukturētu formātu veidā, tad projekta elementu ģeometriskie dati vairumā gadījumu tiek veidoti, izmantojot īpašus CAD rīkus ( 3.1-11. attēls), kas ļauj detalizēti vizualizēt projekta elementus kā līniju kopu (2D) vai ģeometrisko ķermēnu kopu (3D).

3000 BCE - 1960s

physical medium  
(artefact)

1960s to present day

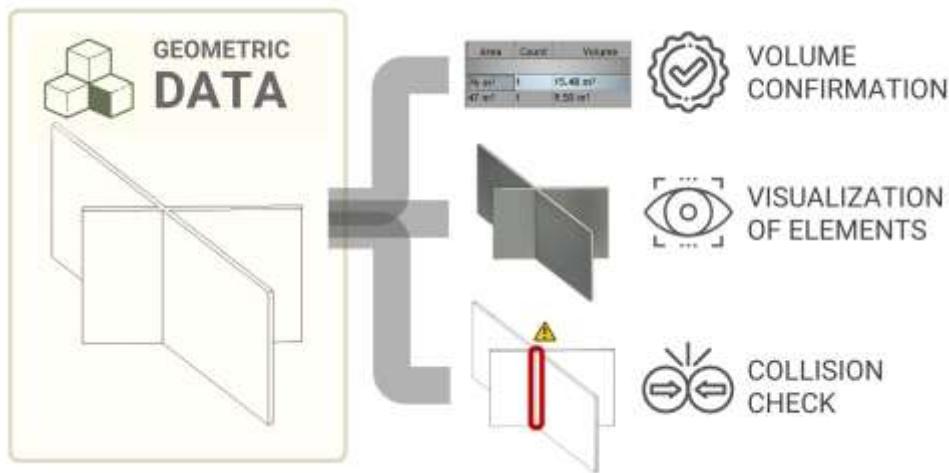
digital project data  
(CAD data)

**Attēls 3.1-11 CAD rīki ir palīdzējuši pārvietot ģeometrisko informāciju no fiziskā datu nesēja uz datu bāzi.**

Strādājot ar ģeometriskiem datiem būvniecībā un arhitektūrā, var izdalīt trīs galvenos ģeometrisko datu lietojumus (3.1-12. attēls):

- **Apjomu apstiprināšana:** ģeometriskie dati, kas tiek generēti CAD programmās (BIM), izmantojot īpašus ģeometriskos kodolus, ir nepieciešami, lai automātiski un precīzi noteiktu projekta elementu apjomus un izmērus. Šie dati ietver automātiski aprēķinātos laukumus, apjomus, garumus un citus svarīgus atribūtus, kas nepieciešami resursu un materiālu plānošanai, budžeta plānošanai un pasūtīšanai
- **Projekta vizualizācija:** projekta izmaiņu gadījumā elementu vizualizācija ļauj automātiski generēt atjauninātus rasējumus dažādās plaknēs. Projekta vizualizācija sākotnējos posmos ļauj panākt ātrāku sapratri starp visiem dalībniekiem, lai taupītu laiku un resursus būvniecības procesa laikā.
- **sadursmju pārbaude:** sarežģītos būvniecības un inženiertehniskajos projektos, kur ir ļoti svarīga vairāku kategoriju elementu (piemēram, cauruļu un sienu) mijiedarbība bez "ģeometriskiem konfliktiem", sadursmju pārbaudei ir ļoti svarīga nozīme. Izmantojot sadursmju noteikšanas programmatūru, varat proaktīvi identificēt iespējamās ģeometriskās pretrunas starp projekta elementiem, novēršot dārgi izmaksājošas klūdas būvniecības procesā.

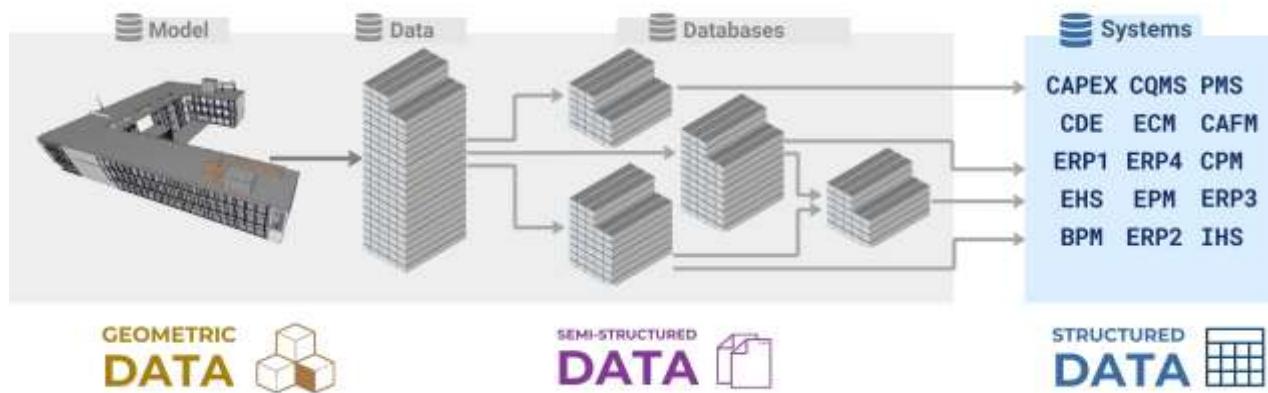
Jau no inženierprojektēšanas biroju pirmsākumiem, sākot ar pirmo sarežģīto konstrukciju būvniecību, būvinženieri sniedza ģeometrisko informāciju rasējumu, līniju un plakano ģeometrisko elementu veidā (uz papirusa, "A0" papīra vai DWG, PDF, PLT formātos), uz kuru pamata meistari un tāmētāji (3.1-11. attēls), pēdējos gadu tūkstošos ar lineālu un transportieru palīdzību apkopoja atribūtu apjomus vai elementu un elementu grupu skaitu. 3.1-11), pēdējos gadu tūkstošos ar lineālu un transportieru palīdzību vāca atribūtu apjomus jeb elementu un elementu grupu daudzumus.



**Attēls 3.1-12 Ģeometrija ir pamats elementu tilpuma parametru iegūšanai, kurus pēc tam izmanto, lai aprēķinātu projekta izmaksas un termiņus.**

Mūsdienās šis manuāli veiktais un laikietilpīgais uzdevums ir pilnībā automatizēts, pateicoties tam, ka mūsdienīgos CAD rīkos (BIM) ir pieejama tilpuma modelēšana, kas ļauj automātiski, izmantojot īpašu ģeometrisko kodolu, iegūt jebkura elementa tilpuma atribūtus bez nepieciešamības manuāli aprēķināt tilpuma parametrus.

Mūsdienīgi CAD rīki ļauj arī klasificēt un iedalīt projekta elementus kategorijās, lai no projekta datubāzes varētu augšupielādēt specifikācijas tabulas izmantošanai dažādās sistēmās, piemēram, izmaksu aprēķina, plānošanas vai CO<sub>2</sub> aprēķina (3.1-13. attēls). Specifikāciju iegūšanu, QTO tabulas un daudzumus, kā arī praktiskus piemērus aplūkosim nodaļā "Daudzumu iegūšana un kvantitatīvais novērtējums".



**Attēls 3.1-13 CAD rīki (BIM) glabā datus datubāzēs, kas ir paredzētas integrācijai un mijiedarbībai ar citām sistēmām.**

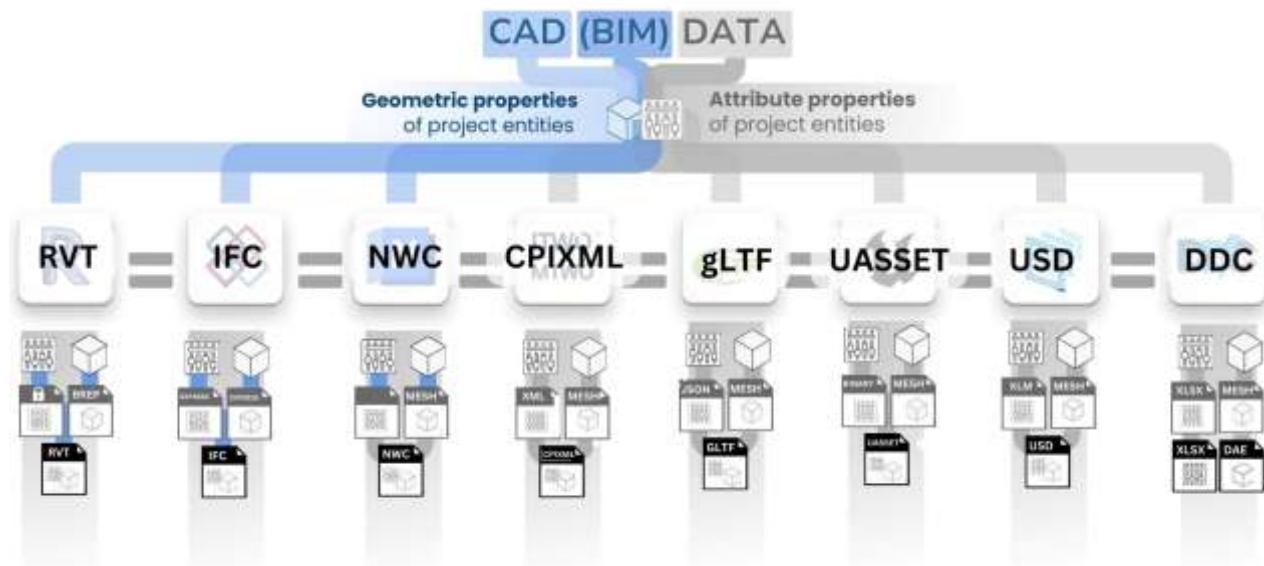
Sakarā ar CAD vidē izmantoto datubāzu un formātu slēgtību, CAD risinājumos radītie ģeometriskie dati faktiski ir kļuvuši par atsevišķu informācijas veidu. Tā apvieno gan elementu ģeometriju, gan metainformāciju (strukturētu vai daļēji strukturētu), kas ietverta specializētos failos un formātos.

## CAD dati: no projektēšanas līdz datu glabāšanai

Mūsdienā CAD un BIM sistēmas glabā datus savos, bieži vien patentētos formātos: DWG, DXF, RVT, DGN, PLN un citos. Šie formāti atbalsta gan 2D, gan 3D objektu attēlojumu, saglabājot ne tikai ģeometriju, bet arī ar objektiem saistītos atribūtus. Šeit ir minēti izplatītākie no šiem formātiem:

- **DWG** ir bināro failu formāts, ko izmanto divdimensiju (un retāk trīsdimensiju) projektēšanas datu un metadatu glabāšanai.
- **DXF** ir teksta formāts 2D un 3D rasējumu apmaiņai starp CAD sistēmām. Tas satur ģeometrijas, slāņu un atribūtu datus, atbalsta gan ASCII, gan bināro attēlojumu.
- **RVT** ir binārais formāts CAD modeļu, tostarp 3D ģeometrijas, elementu atribūtu, attiecību un konstrukcijas parametru glabāšanai.
- **IFC** ir atvērts teksta formāts būvniecības datu apmaiņai starp CAD (BIM) sistēmām. Tas ietver ģeometriju, objektu īpašības un informāciju par to attiecībām.

Papildus šiem formātiem tiek izmantoti arī citi formāti: PLN, DB1, SVF, NWC, CPIXML, BLEND, BX3, USD, XLSX, DAE. Lai gan tie atšķiras pēc mērķa un atvērtības pakāpes (3.1-14. attēls), tie visi var attēlot vienu un to pašu projekta informācijas modeli dažādās formās. Sarežģītos projektos šie formāti bieži tiek izmantoti paralēli, sākot ar rasējumu un beidzot ar projekta modeļu koordinēšanu.



Attēls 3.1-14 Populāri CAD datu glabāšanas formāti apraksta ģeometriju, izmantojot BREP vai MESH parametrus, ko papildina atribūtu datu.

Visi iepriekš minētie formāti ļauj saglabāt datus par katru būvniecības projekta elementu, un visos iepriekš minētajos formātos ir divi galvenie datu veidi:

- **Ģeometriskie parametri** - raksturo objekta formu, atrašanās vietu un izmērus. Ģeometrija un tās izmantošana tiks detalizēti aplūkota grāmatas sestajā daļā, kas veltīta CAD (BIM) risinājumiem;
- **Atribūtu īpašības** - satur dažādu informāciju: materiālus, elementu tipus, tehniskos

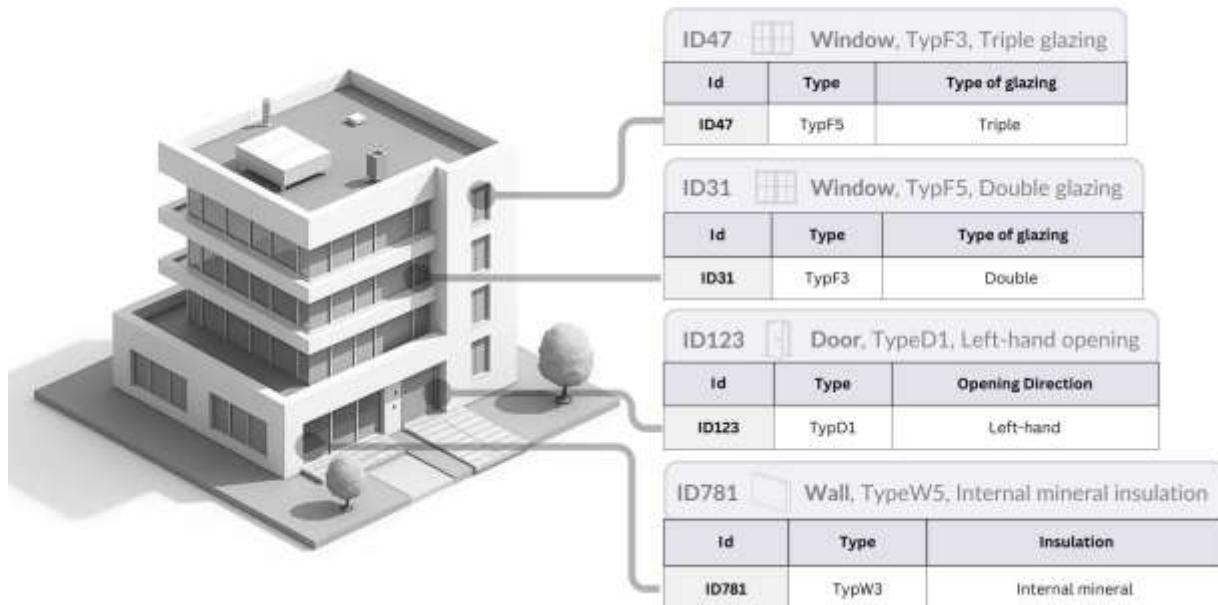
parametrus, unikālos identifikatorus un citas īpašības, kas var būt projekta elementiem.

Atribūtu dati ir īpaši svarīgi mūsdienu projektos, jo tie nosaka objektu ekspluatācijas īpašības, ja jau veikt inženiertehniskos un izmaksu aprēķinus un nodrošina visaptverošu mijiedarbību starp projektēšanas, būvniecības un ekspluatācijas dalībniekiem. Piemēram:

- Logiem un durvīm: konstrukcijas veids, stiklojuma veids, atvēršanas virziens ( 3.2-1. attēls).
- Attiecībā uz sienām tiek reģistrēta informācija par materiāliem, siltumizolāciju un akustiskajām īpašībām.
- Inženiertehniskajām sistēmām tiek saglabāti cauruļvadu, cauruļvadu kanālu, kabeļu trašu un to savienojumu parametri.

Šos parametrus var saglabāt gan pašos CAD-(BIM-) failos, gan ārējās datubāzēs - eksporta, konvertēšanas vai tiešas piekļuves rezultātā iekšējām CAD struktūrām, izmantojot reversās inženierijas rīkus. Šāda pieeja atvieglo projektēšanas informācijas integrāciju ar citām uzņēmuma sistēmām un platformām

**Reversā inženierija** CAD (BIM) kontekstā ir digitālā modeļa iekšējās struktūras ieguves un analīzes process, lai atjaunotu tā loģiku, datu struktūru un atkarības bez piekļuves sākotnējiem algoritmiem vai dokumentācijai.

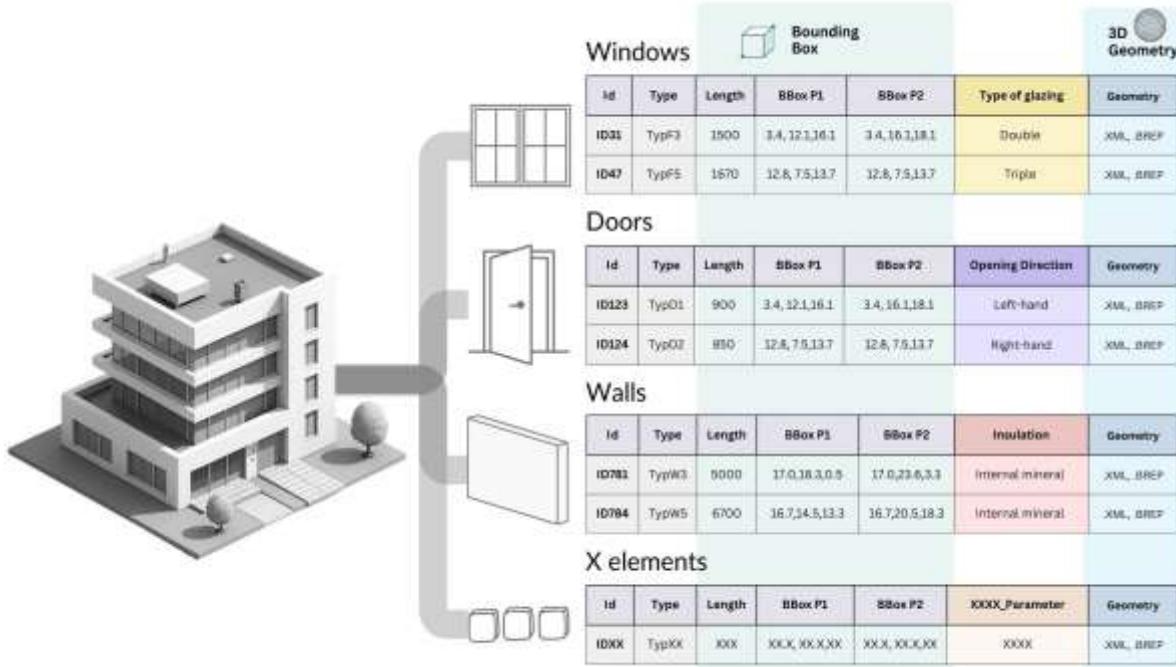


Attēlā 3.1-15. Projekta elements papildus parametriskās vai poligonālās ģeometrijas aprakstam satur informāciju par elementu parametriem un īpašībām.

Rezultātā ap katru elementu tiek izveidots unikāls parametru un īpašību kopums, kas ietver gan katra objekta unikālās īpašības (piemēram, identifikatoru un izmērus), gan elementu grupu kopīgos atribūtus. Tas jau ne tikai analizēt atsevišķus projekta elementus-entitātes, bet arī apvienot tos loģiskās grupās, kuras pēc tam var izmantot citi speciālisti savu uzdevumu veikšanai un aprēķiniem

sistēmās un datubāzēs.

Entitāte ir konkrēts vai abstrakts reālās pasaules objekts, ko var unikāli identificēt, aprakstīt un attēlot datu veidā.

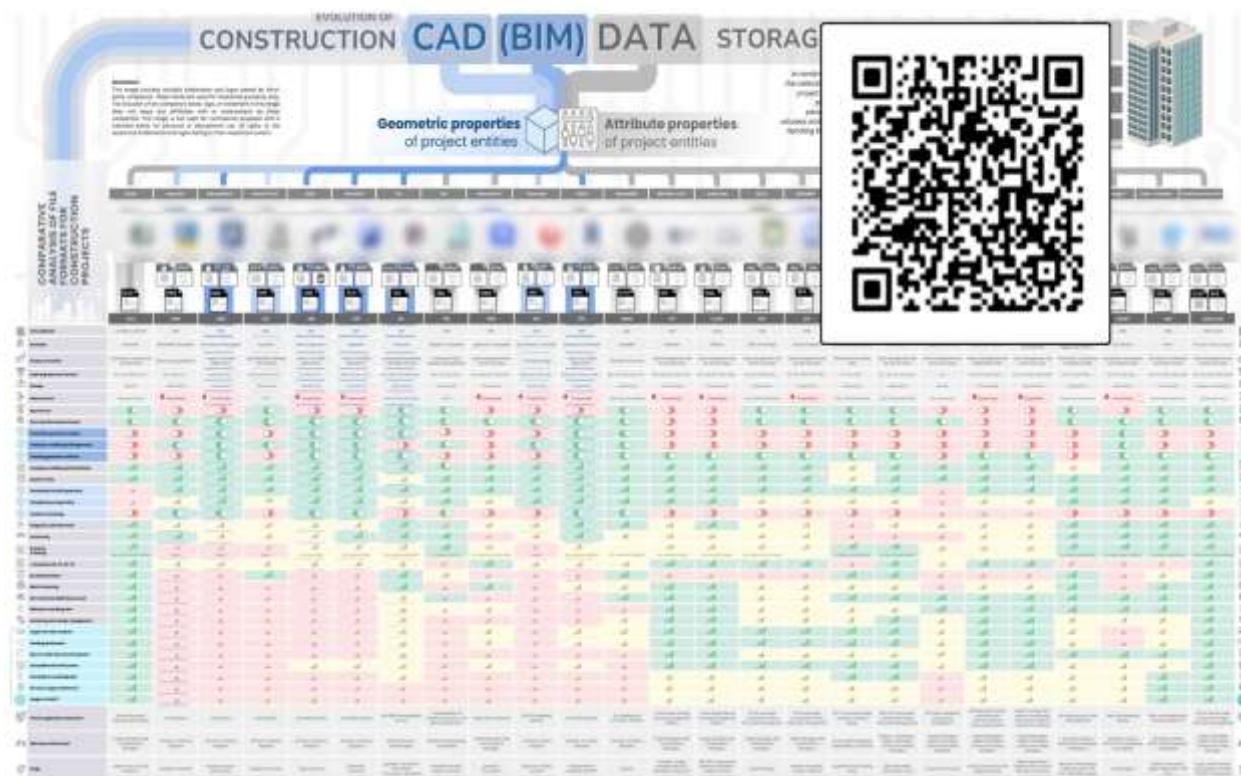


**Attēls 3.1-16** Katrs projekta elements satur atribūtus, kurus vai nu ievada projektētājs, vai aprēķina CAD programmā.

Pēdējo desmitgažu laikā būvniecības nozarē ir izstrādāti daudzi jauni CAD (BIM) formāti, kas vienkāršo datu izveidi, glabāšanu un pārsūtīšanu. Šie formāti var būt slēgti vai atvērti, tabulāri, parametriski vai grafiski. Tomēr to daudzveidība un sadrumstalotība ievērojami sarežģī datu pārvaldību visos projekta dzīves cikla posmos. Informācijas apmaiņai būvniecībā izmantoto galveno formātu salīdzinājuma tabula ir sniepta 3.1-17. attēlā (pilnā versija pieejama ar QR kodu).

Lai risinātu savietojamības un piekļuves CAD datiem problēmas, tiek iesaistīti BIM pārvaldnieki (BIM) un koordinatori, kuru uzdevums ir kontrolēt eksportu, pārbaudīt datu kvalitāti un integrēt CAD (BIM) datu daļas citās sistēmās.

Tomēr, ņemot vērā formātu slēgtību un sarežģītību, šo procesu ir grūti automatizēt, tāpēc speciālistiem daudzas darbības nākas veikt manuāli, bez iespējas izveidot pilnvērtīgus datu apstrādes procesus (pipeline).



The screenshot displays a complex CAD (BIM) software interface. At the top, there are tabs for 'CONSTRUCTION' and 'DATA STORAGE'. Below these, a large grid table represents the core data storage. The columns and rows of the table are filled with various project entity information, including geometric and attribute properties. To the left of the main table, there's a vertical sidebar with sections like 'COMPARATIVE ANALYSIS OF THE PROJECT' and 'PROJECTS'. On the right side, there's a small 3D model of a building and some other graphical elements. A QR code is overlaid on the top right corner of the table area.

**Attēls 3.1-17 Tabula, kurā salīdzināti galvenie datu formāti, kuros tiek glabāta projekta elementu informācija [53].**

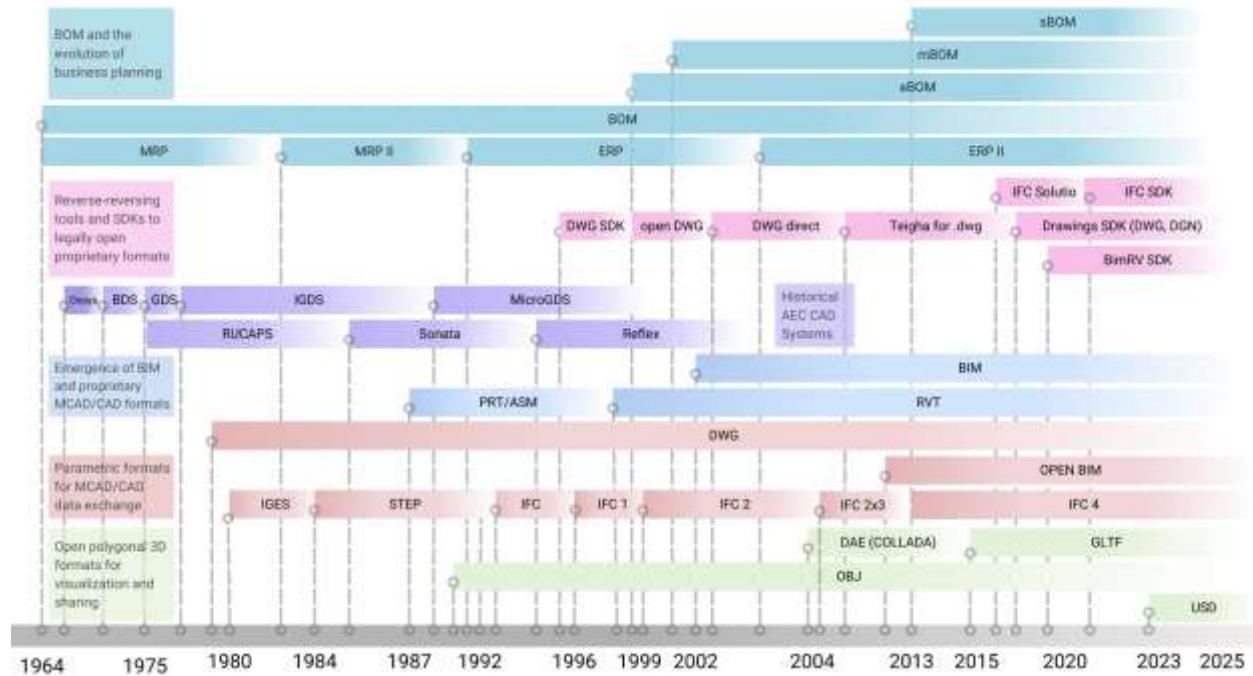
Lai saprastu, kāpēc ir tik daudz dažādu datu formātu un kāpēc lielākā daļa no tiem ir slēgti, ir svarīgi iedziļināties procesos, kas notiek CAD (BIM) programmās, kas tiks detalizēti izpētīti grāmatas sestajā daļā.

Geometrijai pievienoto papildu informācijas slāni CAD sistēmu izstrādātāji ieviesa kā BIM koncepciju (Building Information Modeling) - mārketinga terminu, kas būvniecības nozarē aktīvi tiek popularizēts kopš 2002. gada [54].

### **BIM (BOM) koncepcijas rašanās un CAD izmantošana procesos.**

Būvniecības informācijas modelēšanas (BIM) koncepcija, kas pirmo reizi izklāstīta 2002. gada BIM Baltajā grāmatā [54], radās no CAD programmatūras ražotāju mārketinga iniciatīvām. Tā radās no CAD programmatūras izstrādātāju mārketinga iniciatīvām un bija mēģinājums pielāgot mašīnbūvē jau nostiprinātus principus būvniecības nozares vajadzībām.

BIM iedvesma radās no BOM (Bill of Materials) - materiālu uzskaites saraksta - koncepcijas, kas rūpniecībā tiek plaši izmantota kopš 80. gadu beigām. Mašīnbūvē BOM ļāva sasaistīt CAD sistēmu datus ar PDM (Product Data Management), PLM (Product Lifecycle Management) un ERP sistēmām, nodrošinot holistisku inženiertehniskās informācijas pārvaldību visā produkta dzīves ciklā ( 3.1-8. attēls).



Attēls 3.1-18 Specifikāciju (BOM), informācijas modelēšanas (BIM) un digitālo formātu attīstība inženierbūvniecības nozarē.

Mūsdienu BOM koncepcijas attīstības rezultātā ir radusies paplašināta sistēma - XBOM (Extended BOM), kas ietver ne tikai produkta sastāvu, bet arī uzvedības scenārijus, darbības prasības, ilgtspējas parametrus un datus prognozēšanas analīzei. XBOM būtībā pilda tādu pašu lomu kā BIM būvniecībā: abu pieeju mērķis ir pārvērst digitālo modeli par vienotu patiesības avotu visiem projekta dalībniekiem visā projekta dzīves ciklā.

Būtisks pavērsiens BOM ieviešanā būvniecībā bija pirmā parametriskā CAD (MCAD), kas īpaši pielāgots būvniecības nozarei, ieviešana 2002. gadā. To izstrādāja komanda, kas pirms tam bija radījusi Pro-E® - revolucionāru MCAD sistēmu mašīnbūvei, kas radās 80. gadu beigās un bija kļuvusi par nozares standartu [55].

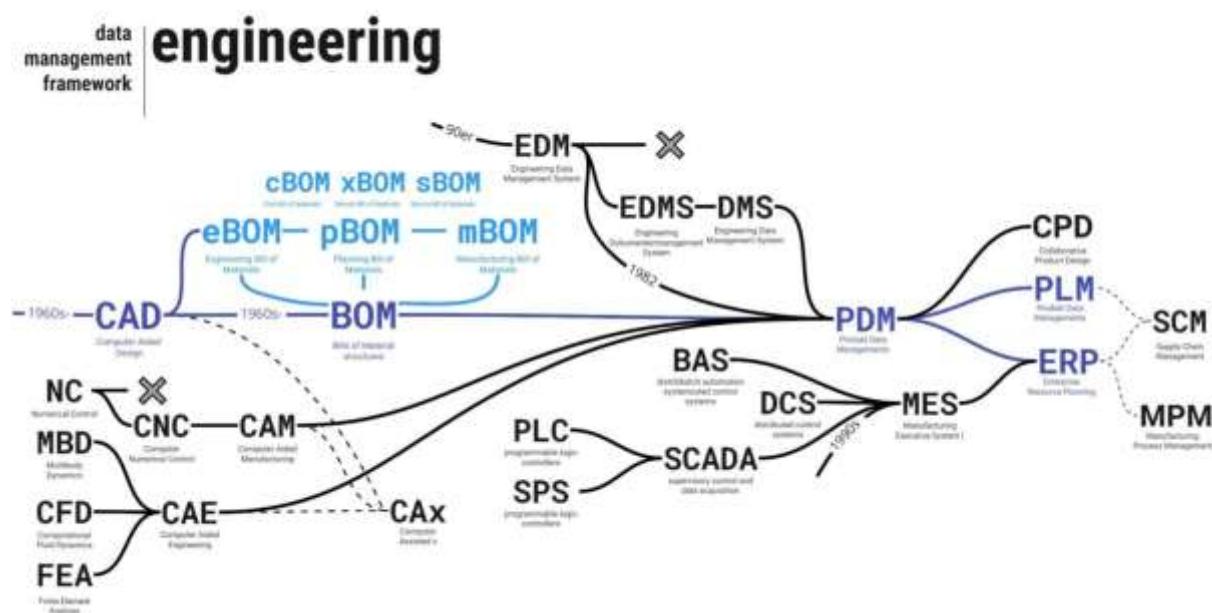
Jau 80. gadu beigās tika izvirzīts mērķis novērst tolaik pastāvošo CAD -programmu ierobežojumus [56]. Galvenais mērķis bija samazināt darbu, kas nepieciešams, lai veiktu izmaiņas konstrukcijas elementu parametros, un dot iespēju atjaunināt modeli, pamatojoties uz datiem ārpus CAD programmām, izmantojot datu bāzi [57]. Svarīgākā loma šajā procesā bija paredzēta parametrizācijai: automātiskai raksturlielumu iegūšanai no datubāzes un to izmantošanai modeļa atjaunināšanai CAD sistēmās.

Pro-E un tā pamatā esošā elementāri parametriskās modelēšanas c BOM koncepcija ir būtiski ietekmējusi CAD - un MCAD - tirgus attīstību [58]. Jau 25 gadus šis modelis darbojas nozarē, un daudzas modernas sistēmas ir kļuvušas par tā konceptuālajiem pēctečiem.

Mērķis ir izveidot sistēmu, kas ir pietiekami elastīga, lai mudinātu inženieri viegli apsvērt dažādus projektus. Un izmaksām, kas rodas, veicot izmaiņas projektā, jābūt pēc iespējas tuvāk nullei. Tradicionālā CAD / CAM programmatūra nereāli ierobežo lētu izmaiņu veikšanu tikai pašā projektēšanas procesa sākumā [59].

- Samuels Heizenbergs, Parametric Technology Corporation® dibinātājs, MCAD -produkta Pro-E izstrādātājs un skolotājs, kurš radījis CAD produktu, izmantojot RVT formātu.

Mašīnbūvē par galvenajām platformām ir kļuvušas PDM, PLM, MRP un ERP sistēmas. Tām ir galvenā loma datu un procesu pārvaldībā, apkopojojot informāciju no CAx sistēmām (CAD, CAM, CAE) un organizējot projektēšanas darbības, pamatojoties uz produkta struktūru (BOM: eBOM, pBOM, mBOM) ( 3.1-18. attēls). Šāda integrācija samazina kļūdu skaitu, novērš datu dublēšanos un nodrošina visaptverošu izsekojamību no projektēšanas līdz ražošanai.



Attēls 3.1-19 Vēsturiski BOM parādījās 60. gados kā veids, kā strukturēt datus no CAx sistēmām un nodot tos vadības sistēmām.

Kad viens no vadošajiem piegādātājiem iegādājās CAD risinājumu, ko izstrādāja kādreizējā Pro-E komanda un kas balstījās uz BOM pieeju, gandrīz nekavējoties tika publicēta BIM Whitepaper sērija (2002-2003)[60][61]. Jau no 2000. gadu vidus būvniecības nozarē sāka aktīvi popularizēt BIM koncepciju, kas ievērojami palielināja interesu par parametrisko programmatūru. Popularitāte pieauga tik strauji, ka mašīnbūves nozares Pro-E - šī ražotāja popularizētā parametriskā CAD - būvniecības dakša faktiski izspieda konkurentus arhitektūras un būvkonstrukciju projektēšanas segmentā (3.1-20. attēls). Līdz 2020. gadu sākumam tas ir de facto nostiprinājis globālo dominanci BIM (CAD) tirgū [62].



Attēls 3.1-20 Google meklēšanas pieprasījumu popularitāte (RVT pret IFC): bijušās Pro-E komandas izveidotā parametriskā CAD ar BOM atbalstu -BIM ir ieguvusi popularitāti gandrīz vairumā pasaules valstu.

Pēdējo 20 gadu laikā saīsinājums BIM ir ieguvis dažādas interpretācijas, kuru daudzās nozīmes saknējas sākotnējās mārketinga koncepcijās, kas radās 2000. gadu sākumā. ISO 19650 standarts, kuram bija liela nozīme termina popularizēšanā, faktiski nodrošināja BIM kā "zinātniski pamatotas" pieejas informācijas pārvaldībai statusu. Tomēr pašā standarta tekstā, kas veltīts datu pārvaldībai visā objektu dzīves ciklā, izmantojot BIM, ir minēts saīsinājums BIM, bet tas nav skaidri definēts

Pārdevēja sākotnējā tīmekļa vietnē, kurā 2002. gadā [60] un 2003. gadā [61] tika publicētas vairākas Baltās grāmatas par BIM, faktiski tika reproducēti mārketinga materiāli par BOM (BMS) un PLM (Product Lifecycle Management) koncepcijām, kas iepriekš tika izmantotas Pro-E mašīnbūves programmatūrā 1990. gados [63].

*Būvniecības informācijas modelēšana - inovatīva, jauna pieeja ēku projektēšanā, būvniecībā un pārvaldībā, ko 2002. gadā ieviesa..... [CAD piegādātāja uzņēmuma nosaukums] - ir mainījusi nozares profesionālu domāšanu visā pasaulei par to, kā tehnoloģijas var izmantot ēku projektēšanā, būvniecībā un pārvaldībā.*

- BIM Baltā grāmata, 2003 [61].

Šajās agrīnajās publikācijās BIM bija tieši saistīts ar centralizētas integrētas datubāzes koncepciju. Kā norādīts 2003. gada Baltajā grāmatā, BIM ir būves informācijas pārvaldība, kurā visi atjauninājumi notiek vienotā repozitorijā, nodrošinot, ka visi rasējumi, griezumi un specifikācijas (BOM - materiālu saraksti) ir sinhronizēti.

*BIM tiek raksturots kā ēkas informācijas pārvaldība, kurā visi atjauninājumi un visas izmaiņas notiek datu bāzē. Tādējādi neatkarīgi no tā, vai jūs strādājat ar shēmām, griezumiem vai lokšņu rasējumiem, viss vienmēr ir saskaņots, konsekvents un aktuāls.*

- CAD uzņēmuma tīmekļa vietne vendor with BIM Whitepaper, 2003 [54].

Ideja par dizaina pārvaldību, izmantojot vienotu integrētu datubāzi, ir plaši apspriesta jau 80. gadu pētījumos. Piemēram, Čārlza Ļstmena (Charles Eastman) BDS koncepcijā [57] bija 43 atsauses uz terminu "datu bāze" (6.1-2. attēls). Līdz 2004. gadam šis skaits bija samazinājies gandrīz uz pus - līdz 23 datiem 2002. gada Baltajā grāmatā par BIM [64]. Un līdz 2000. gadu vidum datubāzu tēma praktiski bija pazudusi no pārdevēju mārketinga materiāliem un digitalizācijas darba kārtības kopumā.

Lai gan sākotnēji BIM sistēmas pamatā bija datu bāze un piekļuve tai, laika gaitā uzsvars tika pārceelts uz ģeometriju, vizualizāciju un 3D. Pats IFC standarta reģistrators 1994. gadā, kurš 2002. gadā publicēja BIM Balto grāmatu, - tas pats piegādātājs - 2000. gadu sākumā Baltajā grāmatā skaidri norādīja uz tādu neitrālu formātu kā IGES, STEP un IFC ierobežojumiem un nepieciešamību pēc tiešas piekļuves CAD datubāzēm:

*Dažādas lietojumprogrammas var būt nesavietojamas, un atkārtoti ievadītie dati var būt neprecīzi [...]. Tradicionālās datorizētās projektēšanas [CAD] rezultāts: lielākas izmaksas, ilgāks laiks līdz produkta laišanai tirgū un zemāka produkta kvalitāte. Mūsdienās visas galvenās lietojumprogrammas izmanto nozares standarta saskarnes zema līmeņa datu apmaiņai. Izmantojot vecos IGES standartus vai jaunos STEP [IFC ir STEP/IGES formāta de facto un de jure kopija], lai apmainītos ar datiem starp dažādu ražotāju lietojumprogrammām, lietotāji var panākt zināmu datu savietojamību starp labākajiem produktiem. Taču IGES un STEP darbojas tikai zemos līmeņos, un ar tiem nevar apmainīties ar tik bagātīgiem datiem kā ar informāciju, ko rada mūsdienu vadošās lietojumprogrammas [...]. Un, lai gan šie un citi standarti tiek uzlaboti gandrīz katru dienu, datu bagātības ziņā tie vienmēr atpaliks no mūsdienu ražotāju produktiem. [...] programmām lietojumprogrammā jāspēj apmainīties un saglabāt datu bagātību, neizmantojot neitrālus tulkotājus, piemēram, IGES, STEP [IFC] vai PATRAN. Tā vietā pamatprogrammām jāspēj tieši piekļūt pamatā esošajai CAD datubāzei, lai netiktu zaudēta informācijas detalizācijas pakāpe un precizitāte.*

- CAD piegādātāju Baltā grāmata (IFC, BIM) "Integrēta projektēšana un ražošana: ieguvumi un pamatojums", 2000 [65].

Tādējādi jau 20. gadsimta 80. gados un 2000. gadu sākumā par galveno digitālās projektēšanas elementu CAD vidē tika uzskatīta datubāze, nevis formāta fails vai neitrāls IFC formāts. Tika ierosināts atteikties no tulkotājiem un nodrošināt lietojumprogrammām tiešu piekļuvi datiem. Tomēr realitātē līdz 20. gadu vidum BIM koncepcija sāka līdzināties stratēģijai "skaldi un valdi", kurā programmatūras ražotāju, kas izmanto slēgtus ģeometriskos kodolus, intereses tiek izvirzītas par prioritāti atvērtās

informācijas apmaiņas attīstības vietā.

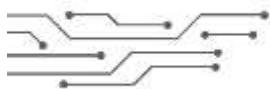
Mūsdienās BIM tiek uztverts kā neatņemama būvniecības nozares sastāvdaļa. Taču pēdējo divdesmit gadu laikā solījumi par vienkāršotu sadarbību un datu integrāciju lielākoties nav īstenoti. Lielākā daļa risinājumu joprojām ir piesaistīti slēgtiem vai neitrāliem formātiem un specializētiem rīkiem. Grāmatas "CAD un BIM: mārketingi, realitāte un projektēšanas datu nākotne būvniecībā" 6. daļā mēs detalizēti aplūkosim BIM vēsturi, atvērto BIM un IFC, kā arī sadarbspējas un ģeometrisko kodolu jautājumus.

Mūsdienās nozare saskaras ar galveno izaicinājumu - pāriet no tradicionālās izpratnes par CAD (BIM) kā modelēšanas rīku uz tā izmantošanu kā pilnvērtīgu datu bāzi. Tas prasa jaunu pieeju darbam ar informāciju, atsakoties no atkarības no slēgtām ekosistēmām un ieviešot atvērtus risinājumus.

Attīstoties reversās inženierijas rīkiem, kas ļauj piekļūt CAD datubāzēm, kā arī izplatoties atvērtā koda un LLM tehnoloģijām, lietotāji un izstrādātāji būvniecības nozarē aizvien vairāk attālinās no programmatūras ražotāju neskaidriem noteikumiem. Tā vietā uzmanība tiek pievērsta tam, kas patiešām ir svarīgi: datiem (datubāzēm) un procesiem.

Aiz moderniem akronīmiem un vizualizācijām slēpjās standarta datu pārvaldības prakse: glabāšana, pārsūtīšana un pārveidošana, t. i., klasiskais ETL process (Extract, Transform, Load). Tāpat kā citās nozarēs, arī būvniecības digitalizācijai ir nepieciešami ne tikai apmaiņas standarti, bet arī skaidri strukturēta heterogēnas informācijas apstrāde.

Lai pilnībā izmantotu CAD (BIM) datu potenciālu, uzņēmumiem ir jāpārdomā sava pieeja informācijas pārvaldībai. Tas neizbēgami novēdīs pie digitālās transformācijas galvenā elementa - to datu unifikācijas, standartizācijas un jēgpilnas struktūrēšanas, ar kuriem būvniecības speciālisti strādā ikdienā.

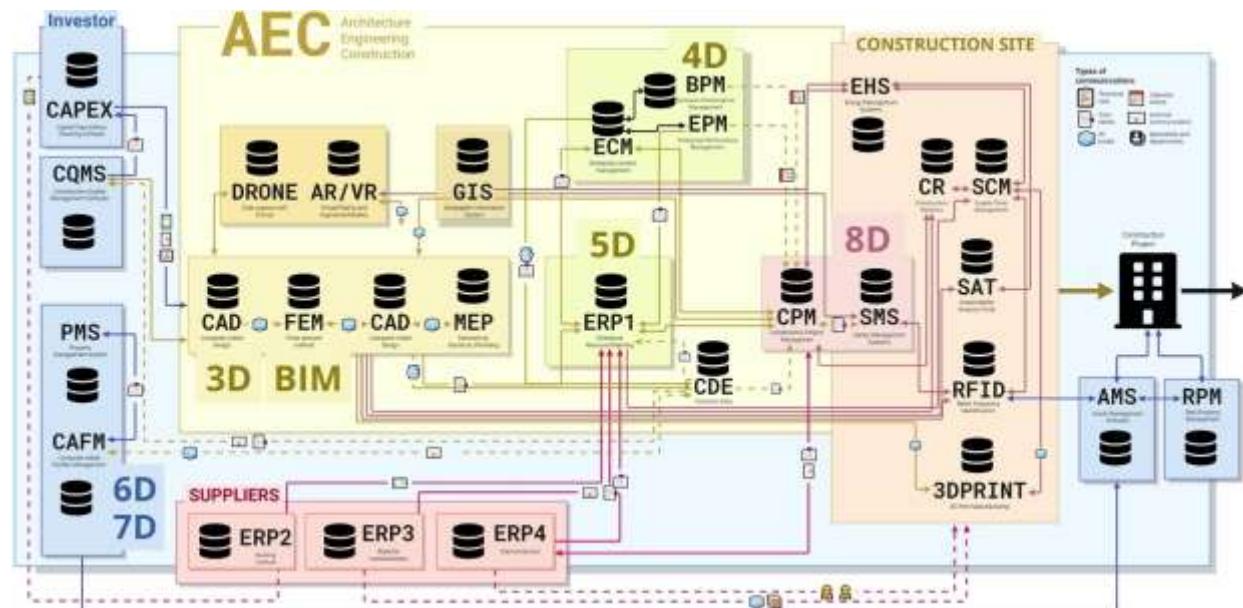


## NODĀĻA 3.2.

# DATU SASKANOŠANA UN STRUKTURĒŠANA

### Sistēmu aizpildīšana ar datiem būvniecības nozarē

Neatkarīgi no tā, vai runa ir par lielām korporācijām vai vidēja lieluma uzņēmumiem, speciālisti ikdienā nodarbojas ar programmatūras sistēmu un datubāzu ar dažādām saskarnēm piepildīšanu ar daudzformātu informāciju (3.2-1. attēls), kam ar vadītāju palīdzību ir savstarpēji jāsadarbojas. Tieši šis savstarpēji mijiedarbojošos sistēmu un procesu komplekss galu galā rada uzņēmuma ienēmumus un peļnu.



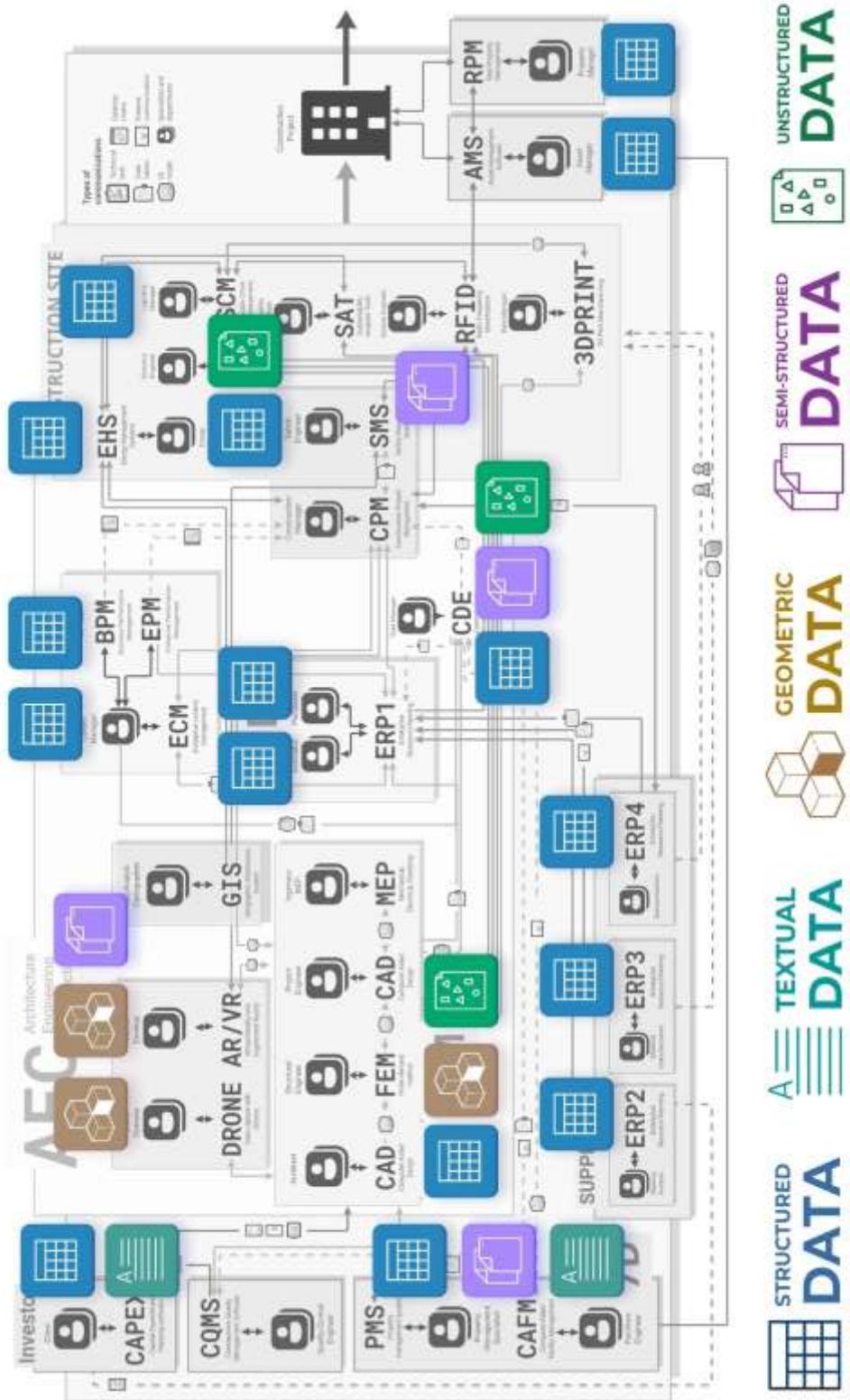
Attēls 3.2-1 Praktiski katras būvniecības nozarē izmantotās sistēmas vai lietojumprogrammas pamatā ir kāda no populārākajām RDBMS datubāzēm.

Katra no iepriekš minētajām un būvniecības nozarē pielietotajām sistēmu kategorijām strādā ar saviem datu tipiem, kas atbilst šo sistēmu funkcionālajai lomai. Lai pārietu no abstraktā līmeņa uz konkrēto, mēs pāriet no datu tipiem uz to attēlojumu formātu un dokumentu veidā.

Iepriekš sniegtajam sistēmu sarakstam (1.2-4. attēls) mēs tagad pievienojam īpašos formātu veidus un dokumentus, ar kuriem tās bieži strādā:

#### **ieguldītājs (CAPEX)**

- Finanšu dati: budžeti, izdevumu prognozes (strukturēti dati).
- Tirdzniecības tendenču dati: tirgus analīze (strukturēti un nestrukturēti dati).
- Juridiskie un līgumiskie dati: līgumi (teksta dati).



3.2-2. attēls Būvniecības nozarē tiek izmantotas daudzas sistēmas ar dažādām saskarnēm, kas strādā ar dažādiem datu veidiem.

### **Vadības sistēmas (PMS, CAFM, CQMS).**

- Projekta dati: grafiki, uzdevumi (strukturēti dati).
- Objekta tehniskās apkopes dati: tehniskās apkopes plāni (teksta un daļēji strukturēti dati).
- Kvalitātes kontroles dati: standarti, pārbaudes ziņojumi (teksta un nestukturēti dati).

### **CAD, FEM un BIM**

- Tehniskie rasējumi: arhitektūras, konstrukciju plāni (ģeometriskie dati, nestukturēti dati).
- Celtniecības modeļi: 3D - modeļi, materiālu dati (ģeometriskie un daļēji strukturētie dati).
- Inženierīniekiskie aprēķini: slodzes analīze (strukturēti dati).

### **Būvlaukuma vadības sistēmas (EHS, SCM)**

- Drošības un veselības dati: drošības protokoli (teksta un strukturēti dati).
- Piegādes ķēdes dati: krājumi, pasūtījumi (strukturēti dati).
- Dienas pārskati: darba laiks, produktivitāte (strukturēti dati).

### **Droni, AR/VR, GIS, 3D - drukāšana.**

- Geodati: topogrāfiskās kartes (ģeometriskie un strukturētie dati).
- Reālā laika dati: video un fotoattēli (nestukturēti dati).
- 3D modeļi - drukāšanai: digitālie rasējumi (ģeometriskie dati).

### **Papildu vadības sistēmas (4D BPM, 5D ERP1).**

- Laika un izmaksu dati: grafiki, tāmes (strukturēti dati).
- Izmaiņu pārvaldība: projekta izmaiņu ieraksti (teksta un strukturēti dati).
- Darbības rezultātu ziņošana: panākumu rādītāji (strukturēti dati).

### **Datu integrācija un saziņa (CDE, RFID, AMS, RPM).**

- Datu apmaiņa: dokumentu apmaiņa, datu modeļi (strukturēti un teksta dati).
- RFID un izsekošanas dati: loģistika, aktīvu pārvaldība (strukturēti dati).
- Uzraudzība un kontrole: sensori objektos (strukturēti un nestukturēti dati).

Tādējādi katra būvniecības nozares sistēma - no būvlaukuma vadības sistēmas līdz operatīvajām datubāzēm - strādā ar sava veida informāciju: strukturētu, teksta, ģeometrisko un citu. "Datu ainava", ar kuru speciālistiem ikdienā nākas strādāt, ir ļoti daudzveidīga. Tomēr vienkāršs formātu uzskaitījums neatklāj reālā darba ar informāciju sarežģītību.

Praksē uzņēmumi saskaras ar to, ka dati, pat ja tie ir iegūti no sistēmām, nav gatavi izmantošanai "tādi, kādi tie ir". Tas jo īpaši attiecas uz tekstiem, attēliem, PDF, CAD failiem un citiem formātiem, kurus ir grūti analizēt ar standarta rīkiem. Tāpēc nākamais svarīgākais solis ir datu transformācija - process, bez kura nevar efektīvi automatizēt apstrādi, analīzi, vizualizāciju un lēmumu pieņemšanu.

## **Datu transformācija: mūsdienu biznesa analīzes kritiskais pamats**

Mūsdienās lielākā daļa uzņēmumu saskaras ar paradoksu: aptuveni 80% ikdienas procesu joprojām balstās uz klasiskiem strukturētiem datiem - pazīstamajām Excel izklājlapām un relāciju datubāzēm (RDBMS) [66]. Tomēr tajā pašā laikā 80% jaunās informācijas, kas nonāk uzņēmumu digitālajā ekosistēmā, ir nestukturēta vai brīvi strukturēta ( 3.2-3. attēls) [52]. Tas ietver tekstu, grafiku,

ģeometriju, attēlus, CAD -modelus, dokumentāciju PDF formātā, audio un video ierakstus, elektronisko saraksti un daudz ko citu.

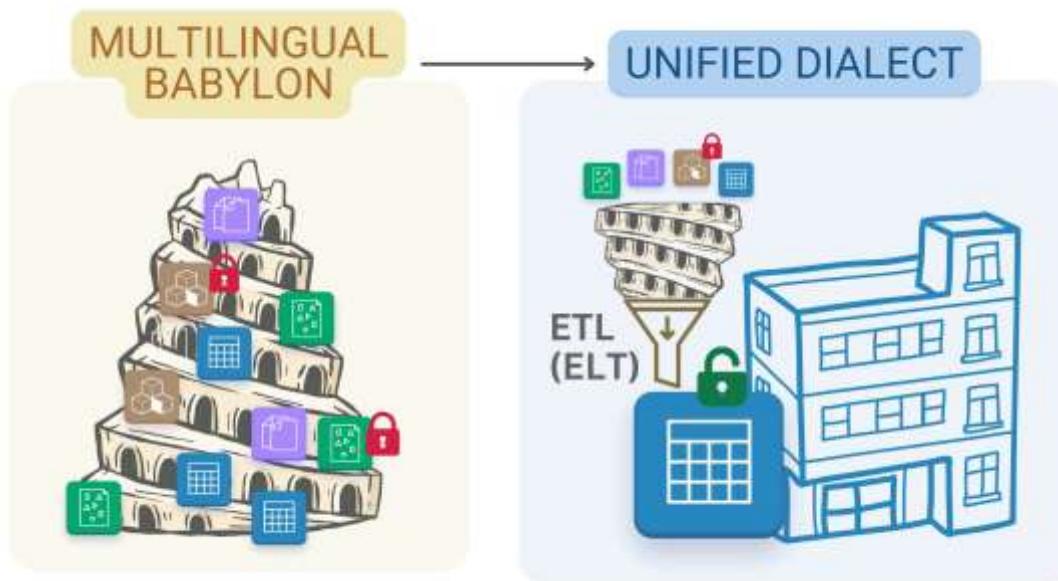
Turklāt nestukturēto datu apjoms turpina strauji pieaugt - tiek lēsts, ka to gada pieaugums ir 55-65% [67]. Šāda dinamika rada nopietnas grūtības integrēt jauno informāciju esošajos uzņēmējdarbības procesos. Šīs daudzformāta datu plūsmas ignorēšana noved pie informācijas nepilnību veidošanās un samazina visas uzņēmuma digitālās vides pārvaldāmību.



**Attēls 3.2-3 Nestukturētu datu ikgadējais pieaugums rada problēmas, lai integrētu straumētu informāciju biznesa procesos.**

Sarežģītu nestukturētu un neskaidru, neskaidri strukturētu datu ignorēšana automatizācijas procesos var radīt būtiskus trūkumus uzņēmuma informācijas ainavā. Mūsdienu pasaulei, kurā informācijas aprite ir nekontrolējama un lavīnveidīga, uzņēmumiem ir jāpieņem hibrīda pieeja datu pārvaldībai, kas ietver efektīvas metodes, kā strādāt ar visu veidu datiem.

Efektīvas datu pārvaldības atslēga ir organizēt, strukturēt un klasificēt dažāda veida datu "Babel" (tostarp nestukturētu, teksta un ģeometrisko formātu, strukturētos vai brīvi strukturētos datos). Šis process pārvērš haotiskas datu kopas organizētās struktūrās, lai tās varētu integrēt sistēmās, tādējādi ļaujot pieņemt uz tām balstītus lēmumus ( 3.2-4. attēls).



**Attēls 3.2-4 Datu pārvaldības nodoļu galvenais uzdevums ir daudzveidīgo un daudzformāta datu "Babilonu" pārvērst strukturētā un kategorizētā sistēmā.**

Viens no galvenajiem šķēršļiem šādai saskanošanai joprojām ir zemais sadarbspējas līmenis starp dažādām digitālajām platformām - "silosiem", par kuriem mēs runājām iepriekšējās nodoļās.

Saskaņā ar ziņojumu Nacionālais standartu un tehnoloģiju institūts (NIST, ASV) uzsver [68], ka nepietiekama datu savietojamība starp dažādām ēku platformām rada informācijas zudumus un ievērojamas papildu izmaksas. Tikai 2002. gadā vien programmatūras savietojamības problēmas izraisīja zaudējumus ASV kapitālbūvniecībā, kas sasniedza 15,8 miljardus ASV dolāru gadā, kur divas trešdaļas no šiem zaudējumiem sedz ēku īpašnieki un apsaimniekotāji, jo īpaši ekspluatācijas un uzturēšanas laikā [68]. Pētījumā arī norādīts, ka datu formātu standartizācija var samazināt šos zaudējumus un uzlabot efektivitāti visā objekta dzīves ciklā.

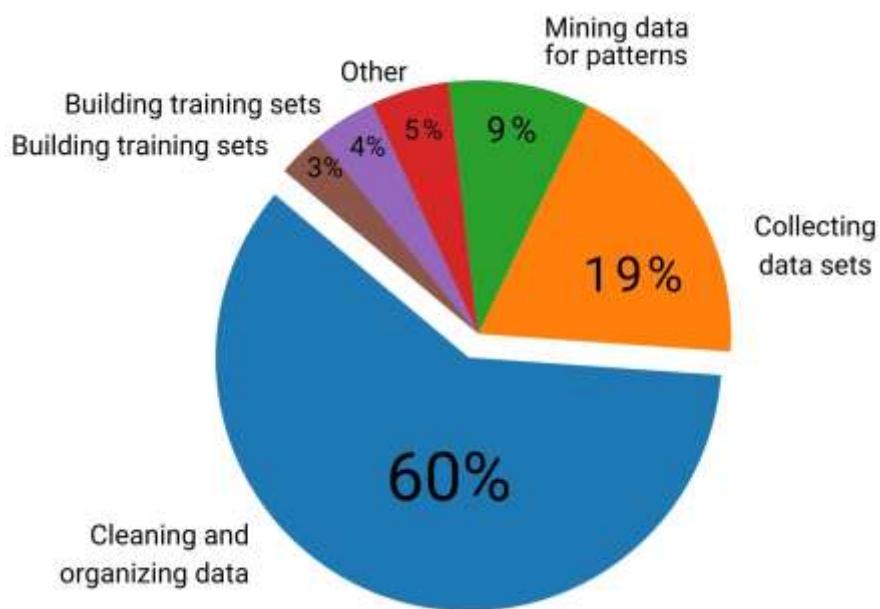
Saskaņā ar 2016. gada CrowdFlower pētījumu [69], kurā piedalījās 16 000 datu zinātnieku visā pasaulē, galvenā problēma joprojām ir "netīri" un daudzformāta dati. Saskaņā ar šo pētījumu vērtīgākais resurss ir nevis galīgās datubāzes vai mašīnmācīšanās modeļi, bet gan laiks, kas veltīts informācijas sagatavošanai.

Tīrīšana, formatēšana un organizēšana aizņem līdz pat 60 procentiem no analītiķa un datu pārvaldnieka laika. Gandrīz piektā daļa tiek tērēta, meklējot un apkopojot vajadzīgās datu kopas, kas bieži vien ir paslēptas datu silososos un nav pieejamas analīzei. Un tikai aptuveni 9 procenti laika tiek tērēti tieši modelēšanai, analīzei, prognožu veidošanai un hipotēžu pārbaudei. Pārējais laiks tiek veltīts komunikācijai, vizualizācijai, pārskatu sagatavošanai un palīgininformācijas avotu izpētei

Vidēji vadītāja darbs ar datiem ir sadalīts šādi ( 3.2-5. attēls):

- **datu tīrīšana un organizēšana (60%):** tīru un strukturētu datu pieejamība var ievērojami samazināt analītiķa darba laiku un paātrināt uzdevumu izpildes procesu.

- **Datu vākšana (19%):** datu zinātnes speciālistiem liela problēma ir atrast atbilstošas datu kopas. Bieži vien uzņēmuma dati ir sakrauti haotiski organizētos "silosos", tādējādi apgrūtinot piekļuvi vajadzīgajai informācijai.
- **Modelēšana/mašīnmācīšanās (9%):** Bieži vien traucē tas, ka klientiem nav skaidru uzņēmējdarbības mērķu. Skaidras misijas formulējuma trūkums var negatīvi ietekmēt pat vislabākā modeļa potenciālu.
- **Citi uzdevumi (5%):** papildus datu apstrādei analītiķiem ir jāveic pētījumi, jāizpēta dati no dažādiem aspektiem, jāpaziņo rezultāti, izmantojot vizualizācijas un pārskatus, un jāiesaka procesu un stratēģiju optimizāciju.



**Attēlā 3.2-5. attēls.Kas ir tas, kam datu pārvaldnieki, kas strādā ar datiem, pavada visvairāk laika (balstoties uz [70]).**

Šīs aplēses apstiprina citi pētījumi. Saskaņā ar 2015. gadā izdevumā BizReport publicēto Xplenty pētījumu [71], no 50% līdz 90% biznesa inteliģences (BI) speciālistu laika (BI) tiek tērēts datu sagatavošanai analīzei.

datu tīrīšana, validēšana un sakārtošana ir būtisks pamats visiem pakārtotajiem datu un analīzes procesiem, kas aizņem līdz pat 90% datu zinātnieku laika.

Šim rūpīgajam darbam, kas galalietotājam nav redzams, ir izšķiroša nozīme. Kļūdas neapstrādātajos datos neizbēgami izkroplo analīzi, ir maldinošas un var radīt dārgas pārvaldības kļūdas. Tāpēc datu attīrišanas un standartizācijas procesi - no dublējošos datu novēršanas un izlaisto datu aizpildīšanas līdz mērvienību saskaņošanai un pielīdzināšanai kopējam modelim - kļūst par mūsdienu digitālās stratēģijas stūrakmeni.

Tādējādi rūpīga datu pārveidošana, attīrišana un standartizācija ne tikai aizņem lielāko daļu speciālistu laika (līdz pat 80% no darba ar datiem), bet arī nosaka to efektīvas izmantošanas iespējas mūsdienu

biznesa procesu ietvaros. Tomēr datu sakārtošana un tīrišana vien ar to neizsmeļ uzdevumu optimāli pārvaldīt uzņēmuma informācijas plūsmas. Organizēšanas un strukturēšanas posmā par to kļūst piemērota datu modeļa izvēle, kas tieši ietekmē darba ar informāciju ērtumu un efektivitāti turpmākajos apstrādes posmos.

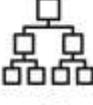
Tā kā datu un uzņēmējdarbības mērķi ir atšķirīgi, ir svarīgi izprast datu modeļu īpašības un spēj izvēlēties vai izveidot pareizo struktūru. Atkarībā no strukturēšanas pakāpes un veida, kādā aprakstītas attiecības starp elementiem, ir trīs galvenie modeļi: strukturēts, brīvi strukturēts un grafisks. Katrs no tiem ir piemērots dažādiem uzdevumiem, un tam ir savas stiprās un vājās puses.

### Datu modeli: attiecības datos un attiecības starp elementiem.

Dati informācijas sistēmās tiek organizēti dažādi - atkarībā no informācijas glabāšanas, apstrādes un pārsūtīšanas uzdevumiem un prasībām. Būtiskākā atšķirība starp datu modeļu veidiem, informācijas uzglabāšanas veidu, ir strukturēšanas pakāpe un veids, kādā tiek aprakstītas attiecības starp elementiem.

Strukturētiem datiem ir skaidra un atkārtojama shēma: tie ir organizēti kā tabulas ar fiksētiem stabījiem. Šāds formāts nodrošina prognozējamību, ērtu apstrādi un efektivitāti, veicot SQL vaicājumus, filtrēšanu un apkopošanu. Piemēri - datu bāzes (RDBMS), Excel, CSV.

Loosely structured data nodrošina elastīgu struktūru: dažādi elementi var saturēt dažādus atribūtu kopumus un tikt saglabāti hierarhiju veidā. Piemēram, JSON, XML vai citi dokumentu formāti. Šādi dati ir ērti, ja nepieciešams modelēt ieliku objektu un attiecības starp tiem, bet, no otras puses, tas sarežģī datu analīzi un standartizāciju ( 3.2-6. attēls).

Data Model	Storage Format	Example
	<b>Relational</b>	CSV, SQL A table of doors in Excel
	<b>Hierarchical</b>	JSON, XML Nested door objects inside a room
	<b>Graph-based</b>	RDF, GraphDB Relationships between building elements

Attēls 3.2-6 Datu modelis ir loģiska struktūra, kas apraksta, kā dati tiek organizēti, uzglabāti un apstrādāti sistēmā.

Atbilstoša formāta izvēle ir atkarīga no mērķiem:

- Ja svarīgs ir filtrēšanas un analīzes ātrums - pietiek ar relāciju tabulām (SQL, CSV, RDBMS,

kolonuveida datubāzēm).

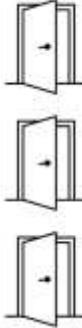
- Ja nepieciešama elastīga struktūra, labāk izmantot JSON vai XML.
- Ja datiem ir sarežģītas attiecības - grafiskās datubāzes nodrošina pārskatāmību un mērogojamību.

Klasiskajās relāciju datubāzēs (RDBMS) katru vienību (piemēram, durvis) attēlo rinda, bet tās īpašības - tabulas kolonas. Piemēram, kategorijas "Durvis" elementu tabulā var būt lauki ID, Augstums, Platums, Ugunsizturība un Telpas ID, kas norāda telpu (3.2-7. attēls).

Klasiskajās relāciju datubāzēs (RDBMS) attiecības veido tabulas, kurās katrs ieraksts ir objekts, bet kolonas - tā parametri. Tabulu formātā dati par durvīm projektā izskatās šādi, kur katra rinda pārstāv atsevišķu elementu - durvis ar to unikālo identifikatoru un atribūtiem, un saikne ar telpu tiek veikta, izmantojot parametru "Room ID".



Door ID	Room ID	Height (mm)	Width (mm)	Fireproof
ID1001	101	2000	900	Yes
ID1002	101	2100	800	No
ID1003	102	2000	850	Yes



**Attēls 3.2-7 Informācija par trim projekta kategorijas "Durvis" elementiem tabulas strukturētā veidā.**

Brīvi strukturētos formātos, piemēram, JSON vai XML, dati tiek glabāti hierarhiskā vai ligzdotā formā, kur elementi var saturēt citus objektus un to struktūra var mainīties. Tas ļauj modelēt sarežģītas attiecības starp elementiem. Līdzīga informācija par durvīm projektā, kas tika reģistrēta strukturētā formā (3.2-7. attēls 3.2-7), ir attēlota brīvi strukturētā formātā (JSON) tādā veidā (3.2-8. attēls 3.2-8. attēls), ka tās klūst par ligzdotiem objektiem telpās (Rooms - ID), kas loģiski atspoguļo hierarhiju.

```

1  {
2    "Rooms": [
3      {
4        "ID": 101,
5        "Doors": [
6          {"ID": 1, "Height": 2000, "Width": 900, "Fireproof": "Yes"},
7          {"ID": 2, "Height": 2100, "Width": 800, "Fireproof": "No"}
8        ]
9      },
10     {
11       "ID": 102,
12       "Doors": [
13         {"ID": 3, "Height": 2000, "Width": 850, "Fireproof": "Yes"}
14       ]
15     }
16   ]
17 }
18

```

Attēls 3.2-8 Informācija par projekta kategorijas "Durvis" elementiem JSON formātā.

Grafa modelī dati ir attēloti kā mezgli (virsotnes) un saites (malas) starp tiem. Tas ļauj vizualizēt sarežģītās attiecības starp objektiem un to atribūtiem. Projekta durvju un telpu datu gadījumā grafiskais attēlojums ir šāds:

- **Mezgli (mezgli)** pārstāv galvenās vienības: telpas (Room 101, Room 102) un durvis (ID1001, ID1002, ID1003).
- **Rievsavienojumi (saites)** parāda attiecības starp šīm vienībām, piemēram, vai durvis pieder pie noteiktas telpas.
- **Atribūti** ir piesaistīti mezgliem un satur vienības īpašības (augstums, platums, durvju ugunsizturība).

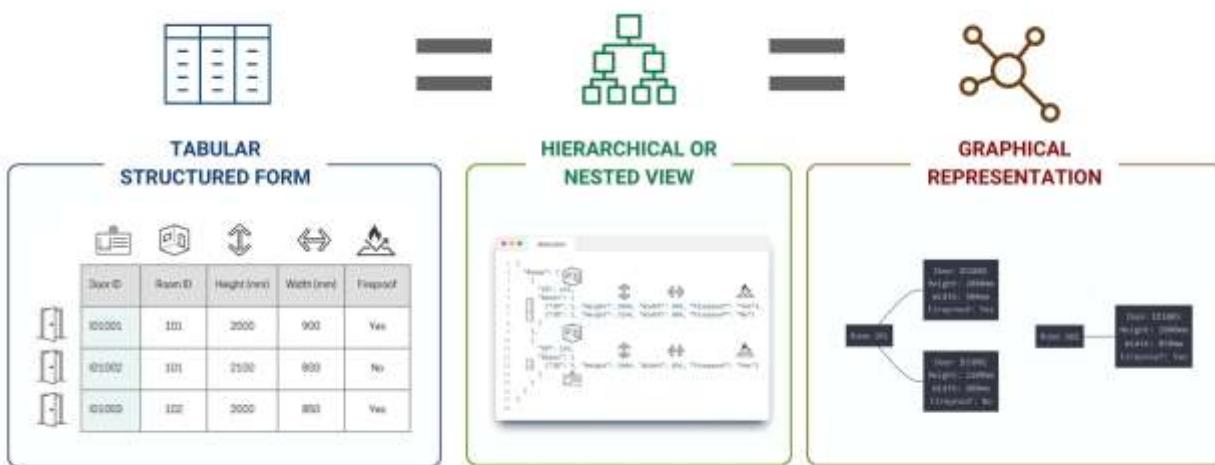


Attēls 3.2-9 Projekta durvju vienības informācija grafika skatā.

Durvju apraksta grafika datu modelī katra telpa un katras durvis ir atsevišķi mezgli. Durvis ar telpām ir saistītas ar malām, kas norāda, ka durvis pieder konkrētai telpai. Durvju atribūti (augstums, platums, ugunsizturība) tiek saglabāti kā attiecīgo mezglu īpašības. Sīkāka informācija par grafa formātiem un to, kā būvniecības nozarē radās grafa semantika, tiks aplūkota nodaļā "semantikas un ontoloģijas rašanās būvniecībā".

Grafu datubāzes ir efektīvas, ja svarīgi ir ne tik daudz paši dati, cik attiecības starp tiem, piemēram, ieteikšanas sistēmās, maršrutēšanas sistēmās vai modelējot sarežģītas attiecības objektu pārvaldības projektos. Grafa formāts atvieglo jaunu attiecību izveidi, ļaujot grafam pievienot jaunus datu tipus, nemainot glabāšanas struktūru. Tomēr, salīdzinot ar relāciju tabulām un strukturetiem formātiem, grafā nav papildu datu savienojamības - pārnesot divdimensiju datubāzes datus grafā, nepalielinās attiecību skaits un nav iespējams iegūt jaunu informāciju.

Datu formai un shēmai jābūt pielāgotai konkrētajam lietošanas gadījumam un veicamajiem uzdevumiem. Lai efektīvi strādātu biznesa procesos, ir svarīgi izmantot tos rīkus un datu modeļus, kas palīdz pēc iespējas ātrāk un vienkāršāk iegūt rezultātus.



Attēls 3.2-10 Vienu un to pašu informāciju par projekta elementiem var saglabāt dažādos formātos, izmantojot dažādus datu modeļus.

Mūsdienās lielākā daļa lielo uzņēmumu saskaras ar pārmērīgas datu sarežģītības problēmu. Katra no simtiem vai tūkstošiem lietojumprogrammu izmanto savu datu modeli, kas rada pārmērīgu sarežģītību - atsevišķs modelis bieži vien ir desmitiem reižu sarežģītāks, nekā nepieciešams, un visu modeļu kopums ir tūkstošiem reižu sarežģītāks. Šī pārmērīgā sarežģītība ievērojami apgrūtina gan izstrādātāju, gan galalietotāju darbu.

Šāda sarežģītība rada nopietnus ierobežojumus uzņēmuma sistēmu izstrādei un uzturēšanai. Katram jaunam elementam modelī ir nepieciešams papildu kods, jaunas logikas ieviešana, rūpīga testēšana un pielāgošana esošajiem risinājumiem. Tas viss palielina izmaksas un palēnina uzņēmuma automatizācijas komandas darbu, pārvēršot pat vienkāršus uzdevumus par dārgiem un laiketielpīgiem

procesiem.

Sarežģītība ietekmē visus datu arhitektūras līmeņus. Relāciju datu bāzēs tā izpaužas kā pieaugošs tabulu un kolonnu skaits, kas bieži vien ir lieki. Objektorientētās sistēmās sarežģītību palielina klašu un savstarpēji saistīto īpašību daudzveidība. Tādos formātos kā XML vai JSON sarežģītība izpaužas kā neskaidras ligzdotas struktūras, unikālas atslēgas un nekonsekventas shēmas.

Datu modeļu pārmērīgā sarežģītība padara sistēmas ne tikai mazāk efektīvas, bet arī apgrūtina to izpratni galalietotājiem un nākotnē - lielajiem valodas modeļiem un LLM aģentiem. Tieši datu modeļu izpratnes un sarežģītības problēma un datu apstrādes sarežģītība rada jautājumu: kā padarīt datus pietiekami vienkāršus lietošanai, lai tie patiesām ātri sāktu būt noderīgi.

Pat tad, ja datu modeļi ir izvēlēti saprātīgi, to lietderība ievērojami samazinās, ja piekļuve datiem ir ierobežota. Patentēti formāti un slēgtas platformas apgrūtina integrāciju, sarežģī automatizāciju un atņem kontroli pār patentētu informāciju, radot ne tikai jaunu datu bunkuru, bet arī aizslēgtu bunkuru, kam var piekļūt tikai ar piegādātāja atļauju. Lai izprastu problēmas mērogu, ir svarīgi apsvērt, kā tieši slēgtās sistēmas ietekmē digitālos procesus būvniecībā.

## **Patentēti formāti un to ietekme uz digitālajiem procesiem**

Viens no galvenajiem izaicinājumiem, ar ko būvniecības uzņēmumi saskaras digitalizācijas laikā, ir ierobežota piekļuve datiem. Tas apgrūtina sistēmu integrāciju, samazina informācijas kvalitāti un apgrūtina efektīvu procesu organizēšanu. Šo grūtību pamatā bieži vien ir patentētu formātu un slēgtu programmatūras risinājumu izmantošana.

Diemžēl līdz šim daudzas būvniecības nozarē izmantotās programmas ļauj lietotājam saglabāt datus tikai patentētos formātos vai mākoņkrātuvēs, kam var piekļūt tikai ar stingri ierobežotu saskarņu palīdzību. Un nereti šie risinājumi ir veidotī, paļaujoties uz vēl noslēgtākām lielāku ražotāju sistēmām. Rezultātā pat tie izstrādātāji, kuri vēlētos piedāvāt atvērtākas arhitektūras, ir spiesti ievērot lielo piegādātāju diktētos noteikumus.

Lai gan modernās būvniecības datu pārvaldības sistēmas aizvien vairāk atbalsta atvērtus formātus un standartus ( 3.1-5. attēls), uz CAD (BIM) balstītas datubāzes un saistītās ERP un CAFM sistēmas joprojām ir izolētas patentētas "salas" nozares digitālajā ainavā ( 3.2-11. attēls).



Attēls 3.2-11 Datu slēgtība un īpašumtiesības rada šķēršļus datu integrācijai un piekļuvei.

Slēgti un monopolizēti formāti un protokoli nav tikai būvniecības nozares problēma. Daudzās tautsaimniecības nozarēs cīņa pret slēgtajiem standartiem un ierobežoto piekļuvi datiem sākās ar inovāciju palēnināšanos ( 3.2-12. attēls), mākslīgu šķēršļu pastāvēšanu jaunu dalībnieku ienākšanai tirgū un pieaugošo atkarību no lielajiem piegādātājiem. Strauji pieaugot datu nozīmei, konkurences iestādēm vienkārši nav laika reaģēt uz jauno digitālo tirgu radītajām problēmām, un rezultātā slēgtie formāti un ierobežotā piekļuve datiem būtībā kļūst par digitālajām "robežām", kas ierobežo informācijas plūsmu un izaugsmi [63].

*Ja mašīnas ražo visu, kas mums nepieciešams, tad mūsu situācija būs atkarīga no tā, kā šīs preces tiek sadalītas. Ikvienas varēs baudīt labklājīgu dzīvi tikai tad, ja mašīnu saražotās bagātības tiks dalītas. Vai arī lielākā daļa cilvēku galu galā dzīvos galējā nabadzībā, ja mašīnu īpašnieki varēs veiksmīgi lobēt pret bagātību pārdali. Pagaidām šķiet, ka situācija attīstās otrajā virzienā, jo tehnoloģijas noved pie arvien lielākas nevienlīdzības [72].*

- Stīvens Hokings, astrofiziķis, 2015 gads

## Monopolies or tight control over critical data formats

**Telecommunications:**  
Proprietary Protocols

1970s-1980s

**Computing Industry:**  
Open Source Movement

1980s

**Document Formats:**  
PDFs and DOCs

Late 1980s to 1990s

**Web Browsing:**  
Browser Wars

Mid-1990s to early 2000s

**Media:**  
Audio and Video Codecs

1990s-2000s



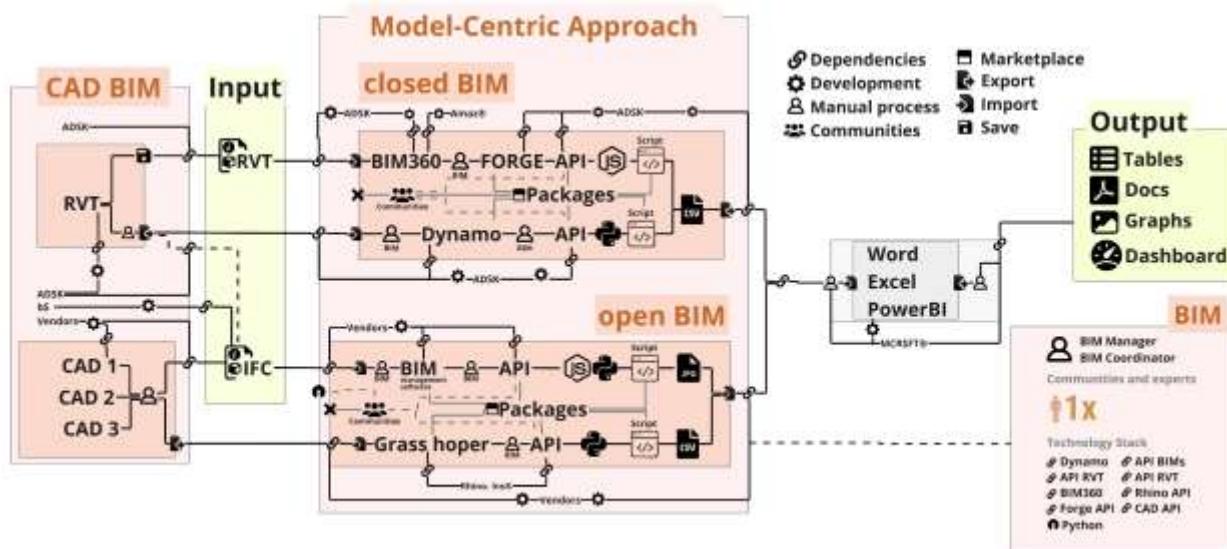
Attēls 3.2-12 Monopola īpašumtiesības uz galvenajiem datu formātiem un protokoliem nav tikai būvniecības nozares problēma.

Tā rezultātā slēgtās piekļuves dēļ datu bāzēm programmām, datu pārvaldnieki, datu analītiķi, IT speciālisti un izstrādātāji, kas veido lietojumprogrammas datu piekļuvei, apstrādei un automatizācijai būvniecības nozarē, šodien saskaras ar daudzām atkarībām no programmatūras piegādātājiem (3.2-13. attēls). Šīs atkarības papildu piekļuves slāņu veidā prasa izstrādāt risinājumus ar specializētiem API -savienojumiem un īpašiem rīkiem un programmatūru.

API (lietojumprogrammu saskarne) ir formalizēta saskarne, ar kuras palīdzību viena programma var mijiedarbīties ar citu, apmainoties ar datiem un funkcijām bez piekļuves pirmkodam. API apraksta, kādus pieprasījumus ārējā sistēma var veikt, kādā formātā tiem jābūt un kādas atbildes tā saņems. Tas ir standartizēts "līgums" starp programmatūras moduļiem.

Lielais atkarību skaits no slēgtiem risinājumiem izraisa to, ka visa koda arhitektūra un uzņēmuma biznesa procesu loģika klūst par "spageti arhitektūru", kas sastāv no rīkiem, kuri ir atkarīgi no programmatūras piegādātāja politikas nodrošināt kvalitatīvu piekļuvi datiem.

Atkarība no slēgtiem risinājumiem un platformām rada ne tikai elastības zudumu, bet arī reālus uzņēmējdarbības riskus. Licencēšanas nosacījumu maiņa, piekļuves slēgšana datiem, formātu vai API struktūras maiņa - tas viss var bloķēt svarīgus procesus. Pēkšņi izrādās, ka vienas tabulas atjaunināšana prasa pārstrādāt veselu integrāciju un savienotāju bloku (3.2-13. attēls), un jebkura liela mēroga programmatūras vai tās API piegādātāja atjaunināšana klūst par potenciālu draudu visas uzņēmuma sistēmas stabilitātei.



Attēls 3.2-13 Piemērs par lielo atkarību skaitu CAD apstrādērada šķēršļus datu integrācijai būvniecības uzņēmumu ekosistēmā.

Izstrādātāji un sistēmu arhitekti šādos apstākļos ir spiesti strādāt nevis, lai sagaidītu, bet lai izdzīvotu. Tā vietā, lai ieviestu jaunus risinājumus, viņi pielāgojas. Tā vietā, lai attīstītu, viņi cenšas saglabāt savietojamību. Tā vietā, lai automatizētu un paātrinātu procesus, viņi tērē laiku nākamo slēgto saskarņu izpētei, API dokumentācijai un bezgalīgai koda pārbūvei.

Darbs ar slēgtiem formātiem un sistēmām ir ne tikai tehnisks izaicinājums - tas ir stratēģisks ierobežojums. Neraugoties uz acīmredzamajām iespējām, ko piedāvā mūsdienu automatizācija, mākslīgais intelekts, LLM un prognozējošā analīze, daudzi uzņēmumi nespēj pilnībā izmantot to potenciālu. Un barjeras, ko rada patentēti formāti (3.2-13. attēls), liedz uzņēmumiem piekļuvi saviem datiem. Iespējams, tā ir būvniecības digitālās transformācijas ironija.

Datu pārredzamība un atvērtas sistēmas nav greznība, bet gan ātruma un efektivitātes priekšnoteikums. Bez atklātības uzņēmēdarbības procesi ir pilni ar nevajadzīgu birokrātiju, daudzslāņainām apstiprināšanas lēdēm un pieaugošu atkarību no HiPPO principa - lēmumu pieņemšana, pamatojoties uz vislabāk apmaksātā cilvēka viedokli.

Tomēr pie apvāršņa ir vērojama paradigmas maiņa. Neraugoties uz patentētu risinājumu dominanci, arvien vairāk uzņēmumu apzinās ceturtās industriālās revolūcijas iedvesmoto arhitektūru ierobežojumus. Šobrīd vektors virzās uz Piektās revolūcijas principiem, kur centrā ir dati kā stratēģisks aktīvs, atvērtas saskarnes (API) un patiesa sistēmu sadarbspēja.

Šī pāreja iezīmē pāreju no slēgtām ekosistēmām uz elastīgām, modulārām digitālajām arhitektūrām, kurās galvenie ir atvērti formāti, standarti un pārredzama datu apmaiņa.

## Atvērtie formāti maina pieeju digitalizācijai

Būvniecības nozare bija viena no pēdējām, kas sāka risināt slēgto un patentēto datu problēmu. Atšķirībā no citām tautsaimniecības nozarēm digitalizācija šeit ir attīstījusies lēni. Tam par iemeslu ir nozares tradicionālais konservatīvais raksturs, atšķirīgu vietējo risinājumu izplatība un dzīļi iesakpojusies papīra dokumentu pārvaldība. Gadu desmitiem galvenie būvniecības procesi balstījās uz fiziskiem rasējumiem, telefona zvaniem un nesinhronizētām datubāzēm. Šajā kontekstā slēgtie formāti ilgu laiku tika uzskatīti par normu, nevis par šķērsli.

Citu nozaru pieredze rāda, ka šķēršļu likvidēšana slēgtajiem datiem veicina inovāciju strauju pieaugumu, paātrina attīstību un palielina konkurenci [73]. Zinātnē atvērto datu apmaiņa ļauj paātrināt atklājumus un attīstīt starptautisko sadarbību. Medicīnā tā var uzlabot diagnostikas un ārstēšanas efektivitāti. Programmatūras inženierijā - radīt koprades un ātras produktu uzlabošanas ekosistēmas.

Saskaņā ar McKinsey ziņojumu "Open Data: Atbloķēt inovācijas un produktivitāti ar informācijas plūsmu" 2013. gadā. [74], atvērtajiem datiem ir potenciāls katru gadu atraisīt 3 līdz 5 triljonus ASV dolāru septiņās galvenajās nozarēs, tostarp būvniecībā, transportā, veselības aprūpē un enerģētikā. Saskaņā ar to pašu pētījumu decentralizētas datu ekosistēmas ļauj lielajiem būvniecības uzņēmumiem un darbuzņēmējiem samazināt programmatūras izstrādes un uzturēšanas izmaksas, paātrinot digitalizācijas ieviešanu.

Pāreja uz atvētajām arhitektūrām, kas jau sen ir sākusies citās ekonomikas nozarēs, pakāpeniski izplatās arī būvniecības nozarē. Lielie uzņēmumi un publiskie klienti, jo īpaši finanšu organizācijas, kas kontrolē ieguldījumus būvniecības projektos, arvien vairāk pieprasīta izmantot atvērtus datus un piekļuvi aprēķinu, aprēķinu un lietojumprogrammu pirmkodam. No izstrādātājiem vairs netiek sagaidīts tikai digitālu risinājumu radīšana un projekta galīgo skaitļu parādišana - no tiem tiek sagaidīta pārredzamība, atkārtojamība un neatkarība no trešo pušu lietojumprogrammu piegādātājiem.

Atvērtā koda risinājumu izmantošana sniedz klientam pārliecību, ka pat tad, ja ārējie izstrādātāji pārtrauks sadarbību vai pametīs projektu, tas neietekmēs iespēju turpināt izstrādāt rīkus un sistēmas. Viens no galvenajiem atvērto datu ieguvumiem ir to spēja novērst lietojumprogrammu izstrādātāju atkarību no konkrētām platformām, lai piekļūtu datiem.

Ja uzņēmums nevar pilnībā atteikties no patentētiem risinājumiem, iespējams kompromiss ir apgrieztās inženierijas metožu izmantošana. Šīs juridiski un tehniski pamatotās metodes ļauj slēgtos formātus pārvērst pieejamākos, strukturētākos un integrēšanai piemērotākos formātos. Tas ir īpaši svarīgi, ja ir nepieciešams pieslēgties mantotām sistēmām vai migrēt informāciju no vienas programmatūras vides uz citu.

Viens no spiltīkajiem piemēriem vēsturē par pāreju uz atvērtiem formātiem un reversās inženierijas (patentētu sistēmu legāla uzlaušana) piemērošanu būvniecībā ir cīņas vēsture par DWG formāta atvēršanu, ko plaši izmanto datorizētās projektēšanas sistēmās (CAD). Reaģējot uz viena programmatūras ražotāja monopolu, 1998. gadā 15 citi CAD ražotāji izveidoja jaunu aliansi "Open DWG", lai nodrošinātu izstrādātājiem bezmaksas un neatkarīgus rīkus darbam ar DWG formātu (faktiskais rasējumu pārsūtīšanas standarts) bez patentētas programmatūras vai slēgtiem API. Šis notikums bija pagrieziena punkts, kas ļāva desmitiem tūkstošu uzņēmumu iegūt brīvu piekļuvi no 80. gadu beigām līdz mūsdienām populārā CAD risinājuma slēgtajam formātam un radīt saderīgus risinājumus, kas veicināja konkurenci CAD tirgū [75]. Šodien "Open DWG" SDK, kas pirmo reizi tika izveidots 1996. gadā, tiek izmantots gandrīz visos risinājumos, kuros iespējams importēt, rediģēt un eksportēt DWG formātu ārpus DWG formāta izstrādātāja oficiālās lietojumprogrammas.

Ari citi tehnoloģiju giganti veic līdzīgas pārmaiņas. Microsoft, kas savulaik simbolizēja patentētu pieejumu, atvēra .NET Framework pirmkodu, sāka izmantot Linux Azure mākoņpakalpojumu infrastruktūrā un iegādājās GitHub, lai nostiprinātu savas pozīcijas atvērtā koda kopienā. [76]. Lai veicinātu inovācijas un sadarbību mākslīgā intelekta aģentu izstrādē, Meta (agrāk Facebook) publiskoja atvērtā pirmkoda mākslīgā intelekta modeļus, piemēram, Llama sēriju. Izpilddirektors Marks Zuckerbergs paredz, ka atvērtā koda platformas nākamajā desmitgadē ienems vadošo lomu tehnoloģiju attīstībā [77].

Atvērtā pirmkoda ir programmatūras izstrādes un izplatīšanas modelis, kurā pirmkods ir atvērts brīvai lietošanai, izpētei, modificēšanai un izplatīšanai.

Atvētie dati un atvērtā koda risinājumi klūst ne tikai par tendenci, bet arī par digitālās ilgtspējības pamatu. Tie nodrošina uzņēmumiem elastību, elastīgumu, elastīgumu, kontroli pār saviem lēmumiem un iespēju paplašināt digitālos procesus, nebūdam iekārti no piegādātāju politikas. Un, kas ir tikpat svarīgi, tie atgriež uzņēmumiem kontroli pār 21. gadsimta vērtīgāko resursu - saviem datiem.

## **Paradigmas maiņa: atvērtā koda kā programmatūras piegādātāju dominances ēras beigas**

Būvniecības nozarē notiek pārmaiņas, kuras nav iespējams novērtēt parastajā veidā. Uz datiem balstītas, uz datiem orientētas pieejas koncepcija un atvērtā koda rīku izmantošana noved pie tā, ka tiek pārdomāti spēles noteikumi, uz kuriem balstās tirgus programmatūras giganti.

Atšķirībā no iepriekšējām tehnoloģiju transformācijām šo pāreju aktīvi nesekmēs pārdevēji. Paradigmas maiņa apdraud to tradicionālos uzņēmējdarbības modeļus, kas balstīti uz licencešanu, abonēšanu un konsultācijām. Jaunā realitāte nav saistīta ar gatavu produktu vai maksas abonementu - tā prasa procesu un domāšanas pārkārtošanu.

Lai pārvaldītu un izstrādātu datu centru risinājumus, kas balstīti uz atvērtām tehnoloģijām, uzņēmumiem būs jāpārdomā iekšējie procesi. Dažādu nodaļu speciālistiem būs ne tikai jāsadarbojas, bet arī jāpārdomā, kā viņi strādā kopā.

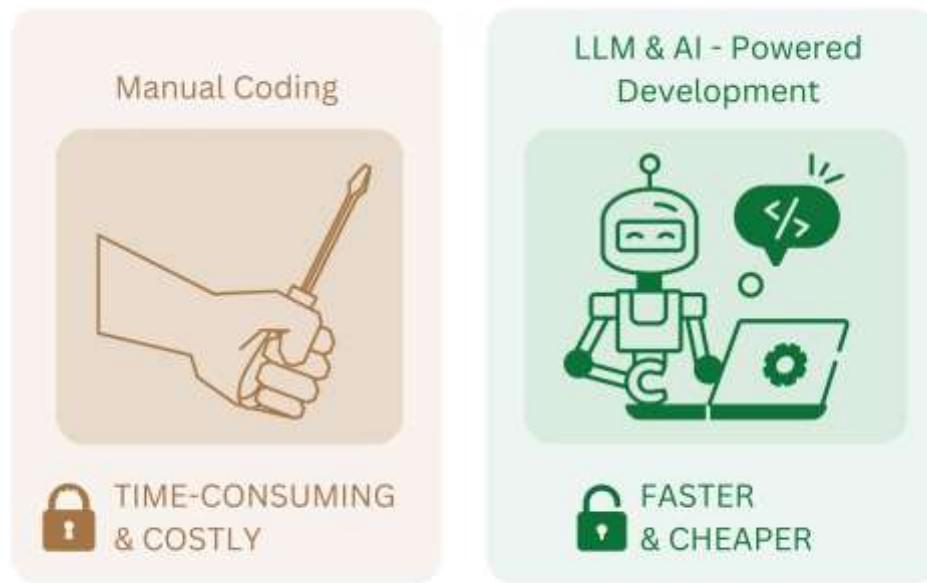
Jaunā paradigma paredz atvērtus datus un atvērtā koda risinājumus, kur īpaša loma programmatūras koda veidošanā būs nevis programmētājiem, bet gan rīkiem, kas balstīti uz mākslīgo intelektu un lielajiem valodas modeļiem (LLM). Jau 2024. gada vidū vairāk nekā 25% no jaunā koda uzņēmumā Google tiks radīti, izmantojot mākslīgo intelektu [78]. Nākotnē kodēšana ar LLM veiks 80% darba tikai 20% laika ( 3.2-14. attēls).

Saskaņā ar McKinsey 2020 pētījumu [79] GPU aizvien biežāk aizstāj CPU analīzē, pateicoties to augstajai veikspējai un moderno atvērtā koda rīku atbalstam. Tas ļauj uzņēmumiem paātrināt datu apstrādi bez ievērojamiem ieguldījumiem dārgā programmatūrā vai nepietiekamu speciālistu algošanas.

Tādi vadošie konsultāciju uzņēmumi kā McKinsey, PwC un Deloitte uzsver, ka arvien pieaug atvērto standartu un atvērtā koda lietojumprogrammu nozīme visās nozarēs.

Saskaņā ar PwC Open Source Monitor 2019 ziņojumu [80] 69% uzņēmumu ar 100 vai vairāk darbiniekiem apzināti izmanto atvērtā koda risinājumus. Īpaši aktīvi OSS tiek izmantots lielos uzņēmumos: Uzņēmumos ar 200-499 darbiniekiem to izmanto 71%, 500-1999 darbinieku kategorijā - 78%, bet uzņēmumos ar vairāk nekā 2000 darbiniekiem - līdz pat 86%. Saskaņā ar Synopsys OSSRA 2023 ziņojumu 96% analizēto kodu bāzu saturēja atvērtā pirkoda komponentes [81].

Nākotnē izstrādātāja uzdevums nebūs manuāli rakstīt kodu, bet gan izstrādāt datu modeļus, plūsmas arhitektūras un pārvaldīt mākslīgā intelekta aģentus, kas pēc pieprasījuma izveido pareizos aprēķinus. Lietotāja saskarnes kļūs minimālistiskas, un mijiedarbība būs balstīta uz dialogiem. Klasiskā programmēšana aizstās vietu augsta līmeņa digitālo risinājumu projektēšanai un orķestrēšanai ( 3.2-14. attēls). Pašreizējās tendencies, piemēram, zema koda platformas ( 7.4-6. attēls) un LLM atbalstītas ekosistēmas ( 7.4-4. attēls), ievērojami samazinās IT sistēmu izstrādes un uzturēšanas izmaksas.



**Attēls 3.2-14 Lai gan šodien lietojumprogrammas manuāli veido programmētāji, nākotnē ievērojamu daļu koda ģenerēs mākslīgais intelekts un uz LLM balstīti risinājumi.**

Šī pāreja atšķirsies no iepriekšējām, un maz ticams, ka lielie programmatūras ražotāji būs tās katalizatori.

Hārvarda Biznesa skolas pētījumā "Atvērtā pirmkoda programmatūras vērtība" 2024 [40] atvērtā pirmkoda programmatūras kopējā vērtība tiek lēsta no diviem aspektiem. No vienas pusēs, ja mēs aprēķinātu, cik daudz būtu nepieciešams, lai visus esošos atvērtā koda risinājumus izveidotu no nulles, summa būtu aptuveni 4,15 miljardi ASV dolāru. No otras pusēs, ja mēs iedomājamies, ka katrs uzņēmums pats izstrādā savus atvērtā pirmkoda risinājumu analogus (kas notiek visur), bez piekļuves esošajiem rīkiem, tad kopējās uzņēmējdarbības izmaksas sasniegta kolosālos 8,8 triljonus dolāru - tā ir pieprasījuma cena.

Nav grūti nojaust, ka neviens no lielākajiem programmatūras piegādātājiem nav ieinteresēts samazināt programmatūras tirgu, kura potenciālā vērtība ir 8,8 triljoni ASV dolāru, līdz tikai 4,15 miljardiem ASV dolāru. Tas nozīmētu samazināt pieprasījuma apjomu vairāk nekā 2000 reižu. Šāda transformācija ir vienkārši nerentabla piegādātājiem, kuru uzņēmējdarbības modeļi ir balstīti uz gadiem ilgušu klientu atkarības uzturēšanu no slēgtiem risinājumiem. Tāpēc uzņēmumi, kas sagaida, ka kāds tiem piedāvās ērtu un atvērtu gatavu risinājumu, var būt vīlušies - šie pārdevēji vienkārši neparādīsies.

Pāreja uz atvērtu digitālo arhitektūru nenozīmē darbavietu vai ieņēmumu zaudēšanu. Gluži pretēji, tā rada priekšnoteikumus elastīgiem un adaptīviem uzņēmējdarbības modeļiem, kas galu galā var izspiest tradicionālo licenču un iepakotas programmatūras tirgu.

Licenču pārdošanas vietā - pakalpojumi, slēgtu formātu vietā - atvērtas platformas, atkarības no

piegādātāja vietā - neatkarība un iespēja veidot risinājumus reālām vajadzībām. Tie, kas agrāk vienkārši izmantoja rīkus, varēs kļūt par to līdzautoriem. Un tie, kas spēj strādāt ar datiem, modeļiem, scenārijiem un loģiku, nonāks nozares jaunās digitālās ekonomikas centrā. Plašāk par šīm pārmaiņām un par to, kādas jaunas lomas, uzņēmējdarbības modeļi un sadarbības formāti veidojas saistībā ar atvērtajiem datiem, mēs runāsim grāmatas pēdējā, desmitajā, daļā.

Uz atvērtiem datiem un atvērtu kodu balstīti risinājumi jaus uzņēmumiem koncentrēties uz uzņēmējdarbības procesu efektivitāti, nevis cīnīties ar novecojušām API un integrēt slēgtas sistēmas. Apzināta pāreja uz atvērto arhitektūru var ievērojami uzlabot produktivitāti un samazināt atkarību no piegādātājiem.

Pāreja uz jauno realitāti nozīmē ne tikai programmatūras izstrādes pieejas maiņu, bet arī paša darba ar datiem principa pārdomāšanu. Šo pārmaiņu centrā ir nevis kods, bet gan informācija - tās struktūra, pieejamība un interpretējamība. Un tieši šeit priekšplānā izvirzās atvērti un strukturēti dati, kas kļūst par jaunās digitālās arhitektūras neatņemamu sastāvdaļu.

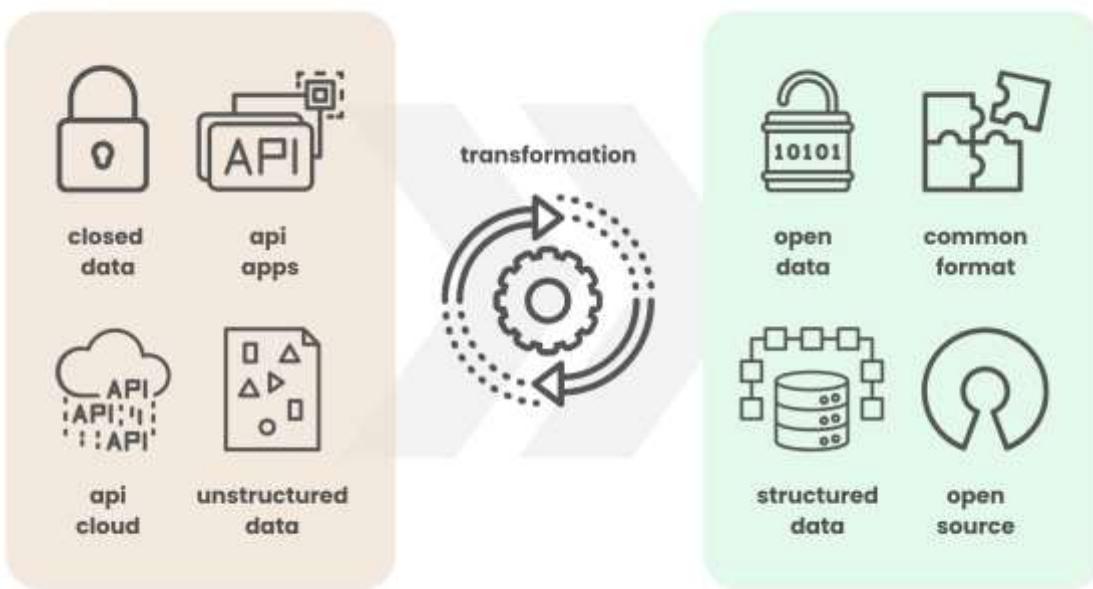
## **Strukturēti atvērti dati: digitālās transformācijas pamats**

Ja iepriekšējās desmitgadēs uzņēmējdarbības ilgtspēju lielā mērā noteica programmatūras risinājumu izvēle un atkarība no konkrētiem piegādātājiem, tad mūsdienu digitālajā ekonomikā galvenais faktors ir datu kvalitāte un spēja ar tiem efektīvi strādāt. Atvērtā koda kods ir svarīga jaunās tehnoloģiskās paradigmas daļa, taču tā potenciāls ir patiesi atraisīts tikai tad, ja dati ir saprotami, organizēti un mašīnlasāmi. Starp visu veidu datu modeļiem strukturēti atvērti dati kļūst par ilgtspējīgas digitālās transformācijas stūrakmeni.

Galvenā strukturētu atvērto datu priekšrocība ir viennozīmīga interpretācija un automatizētas apstrādes iespēja. Tas jauj ievērojami paaugstināt efektivitāti gan atsevišķu darbību līmenī, gan visā organizācijā.

Saskaņā ar Deloitte ziņojumu "Datu pārneses process uzņēmumu pārveidē". [82], sadarbība ar IT, lai pārvaldītu strukturētu datu pārnesi, ir joti svarīga. Saskaņā ar Apvienotās Karalistes valdības ziņojumu "Datu analīze un mākslīgais intelekts valdības projektu īstenošanā" (2024) [83], lai uzlabotu projektu pārvaldības efektivitāti, ir svarīgi novērst šķēršļus datu apmaiņai starp dažādiem projektiem un organizācijām. Dokumentā uzsvērts, ka datu formātu standartizācija un atvērto datu principu ieviešana var novērst informācijas dublēšanos, samazināt laika zudumus un uzlabot prognožu precizitāti.

Būvniecības nozarē, kur tradicionāli dominē liela sadrumstalotība un formātu daudzveidība, strukturētam apvienošanas procesam un strukturētiem atvērtiem datiem ir izšķiroša nozīme saskaņotu un pārvaldāmu procesu veidošanā ( 4.1-14. attēls). Tie jauj projekta daļībniekiem koncentrēties uz produktivitātes uzlabošanu, nevis uz tehnisko problēmu risināšanu, kas saistītas ar slēgto platformu, datu modeļu un formātu nesaderību.



Attēlā: 3.2-15 Atvērtie strukturētie dati samazina atkarību no programmatūras risinājumiem un platformām un paātrina inovāciju.

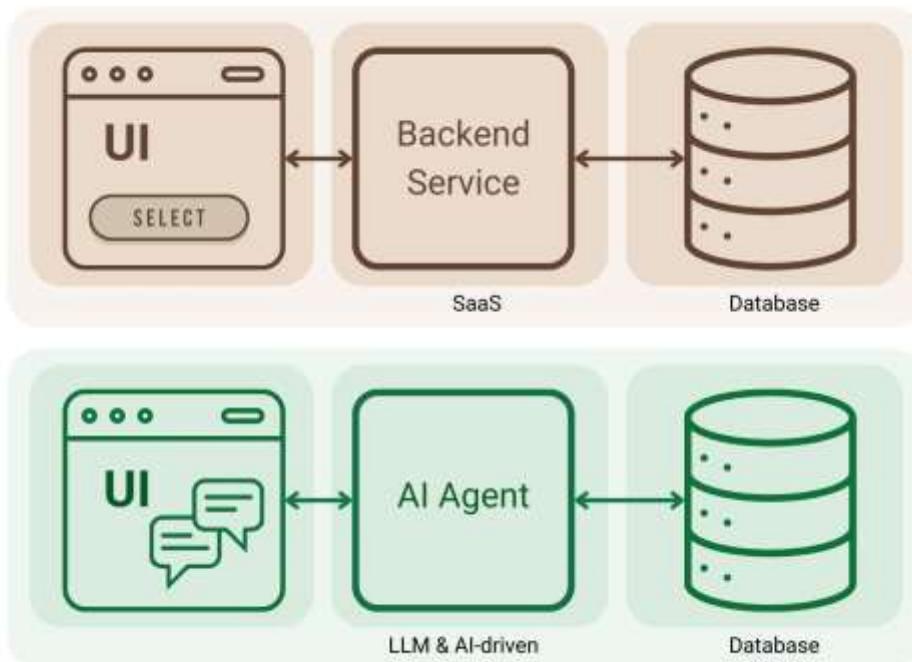
Mūsdienu tehnoloģiju rīki, par kuriem sīkāk runāsim tālāk grāmatā, ļauj ne tikai apkopot informāciju, bet arī automātiski to attīrīt: novērst dublēšanos, labot kļūdas un normalizēt vērtības. Tas nozīmē, ka analītiķi un inženieri strādā nevis ar atšķirīgiem dokumentiem, bet gan ar sakārtotu zināšanu bāzi, kas piemērota analīzei, automatizācijai un lēmumu pieņemšanai.

*Padariet to pēc iespējas vienkāršāku, bet ne vienkāršāku.*

- Alberts Einšteins, teorētiskais fiziķis (citāta autorība ir apstrīdēta [84]).

Mūsdienās lielāko daļu lietotāja saskarņu darbam ar datiem var izveidot automātiski - bez nepieciešamības manuāli rakstīt kodu katram uzņēmējdarbības gadījumam. Tam nepieciešams infrastruktūras slānis, kas saprot datu struktūru, modeli un loģiku bez papildu instrukcijām (4.1-15. attēls). Tieši strukturēti dati padara šo pieeju iespējamu: veidlapas, tabulas, filtrs un skatus var automātiski generēt ar minimālu programmēšanas darbu.

Svarīgākās, lietotājam kritiski svarīgās saskarnes joprojām var būt jāpilnveido manuāli. Taču vairumā gadījumu - un tas ir no 50 līdz 90 procentiem darba scenāriju - pietiek ar automātisku lietojumprogrammu un aprēķinu generēšanu, neizmantojot šim nolūkam paredzētas speciālas lietojumprogrammas (3.2-16. attēls), kas ievērojami samazina izstrādes un uzturēšanas izmaksas, samazina kļūdu skaitu un paātrina digitālo risinājumu ieviešanu.



**Attēls 3.2-16 Arhitektūras modeļi darbam ar datiem: tradicionālā lietojumprogrammu arhitektūra un uz mākslīgo intelekta orientēts modelis ar LLM.**

Pāreja no arhitektūrām, kas balstītas uz individuālām lietojumprogrammām, uz inteliģenti pārvaldītām sistēmām, kuru pamatā ir valodas modeļi (LLM), ir nākamais digitālās evolūcijas solis. Šādā arhitektūrā strukturēti dati kļūst ne tikai par glabāšanas objektu, bet arī par pamatu mijiedarbībai ar mākslīgā intelekta rīkiem, kas spēj analizēt, interpretēt un ieteikt darbības, pamatojoties uz kontekstu.

Turpmākajās nodaļās aplūkosim reālus piemērus, kā īstenot uz atvērtiem struktūrētiem datiem balstītu arhitektūru, un parādīsim, kā valodas modeļi tiek izmantoti, lai automātiski interpretētu, validētu un apstrādātu datus. Šie praktiskie gadījumi palīdzēs jums labāk izprast, kā jaunā digitālā loģika darbojas darbībā - un kādas priekšrocības tā sniedz uzņēmumiem, kas ir gatavi transformācijai.



## NODĀLA 3.3.

### **LLM UN TO LOMA DATU APSTRĀDĒ UN UZNĒMĒJDARBĪBAS PROCESOS**

**LLM tērzēšanas istabas: ChatGPT, LlaMa, Mistral, Claude, DeepSeek, QWEN, Grok datu procesu automatizēšanai.**

Lielo valodas modeļu (LLM) rašanās bija dabisks turpinājums virzībai uz strukturētiem atvērtiem datiem un atvērtā koda filozofijai. Kad dati kļūst organizēti, pieejami un mašīnlasāmi, nākamais solis ir rīks, kas var mijiedarboties ar šo informāciju bez nepieciešamības rakstīt sarežģītu kodu vai speciālām tehniskām zināšanām.

Mūžizglītības programmas ir tiešs atvērtības produkts: lielas atvērtas datu kopas, publikācijas un atvērtā koda kustība. Bez atvērtiem zinātniskajiem raksti, publiiski pieejamiem teksta datiem un sadarbības attīstības kultūras nebūtu ne ChatGPT, ne citu LLM. LLM savā ziņā ir cilvēces uzkrāto digitālo zināšanu "destilāts", kas apkopots un izglītojams, izmantojot atklātības principus.

Mūsdieni lielie valodas modeļi (LLM - Large Language Models), piemēram, ChatGPT ® (OpenAI), LlaMa ™ (Meta AI), Mistral DeepSeek™, Grok ™ (xAI), Claude ™ (Anthropic), QWEN™ nodrošina lietotājiem iespēju formulēt pieprasījumus datiem dabiskajā valodā. Tas padara darbu ar informāciju pieejamu ne tikai izstrādātājiem, bet arī analītiķiem, inženieriem, plānotājiem, vadītājiem un citiem speciālistiem, kas iepriekš bija attālināti no programmēšanas

**LLM** (Large Language Model) ir mākslīgais intelekts, kas ir apmācis saprast un ģenerēt tekstu, pamatojoties uz milzīgu datu daudzumu, kas savākti no visa interneta. Tas spēj analizēt kontekstu, atbildēt uz jautājumiem, iesaistīties dialogā, rakstīt tekstus un ģenerēt programmatūras kodu.

Ja agrāk datu vizualizācijai, apstrādei vai analīzei bija nepieciešamas speciālas programmēšanas valodas zināšanas: Python, SQL, R vai Scala, kā arī prasme strādāt ar tādām bibliotēkām kā Pandas, Polars vai DuckDB un daudzām citām, sākot ar 2023. gadu situācija ir radikāli mainījusies. Tagad lietotājs var vienkārši aprakstīt, ko viņš vēlas iegūt, un modelis pats ģenerēs kodu, izpildīs to, parādīs tabulu vai grafiku un paskaidros rezultātu. Pirmo reizi gadu desmitu laikā tehnoloģiju attīstība ir gājusi nevis sarežģīšanas, bet gan radikālas vienkāršošanas un pieejamības ceļu.

Šis princips - "apstrādāt datus ar vārdiem (ieteikumiem)". - iezīmēja jaunu posmu darba ar informāciju evolūcijā, efektīvi pārceļot risinājumu radīšanu vēl augstākā abstrakcijas līmenī. Tāpat kā kādreiz lietotājiem vairs nebija nepieciešams izprast interneta tehniskos pamatus, lai darbinātu interneta veikalus vai veidotu tīmekļa vietnes, izmantojot WordPress, Joomla un citas atvērtā koda moduļu sistēmas (av tor grāmata ar šādām sistēmām strādā kopš 2005. gada, tostarp izglītības un inženiertehniskās tiešsaistes platformas). - Tas savukārt ir izraisījis digitālā satura un tiešsaistes uzņēmējdarbības uzplaukumu - šodien inženieri, analītiķi un vadītāji var automatizēt darba procesus

bez programmēšanas valodu zināšanām. To atvieglo jaudīgi LLM - gan bezmaksas, gan atvērtā pirmkoda, piemēram, LLaMA, Mistral, Qwen, DeepSeek un citi -, kas progresīvas tehnoloģijas padara pieejamas visplašākai auditorijai.

## **Lieli LLM valodas modeļi: kā tas darbojas**

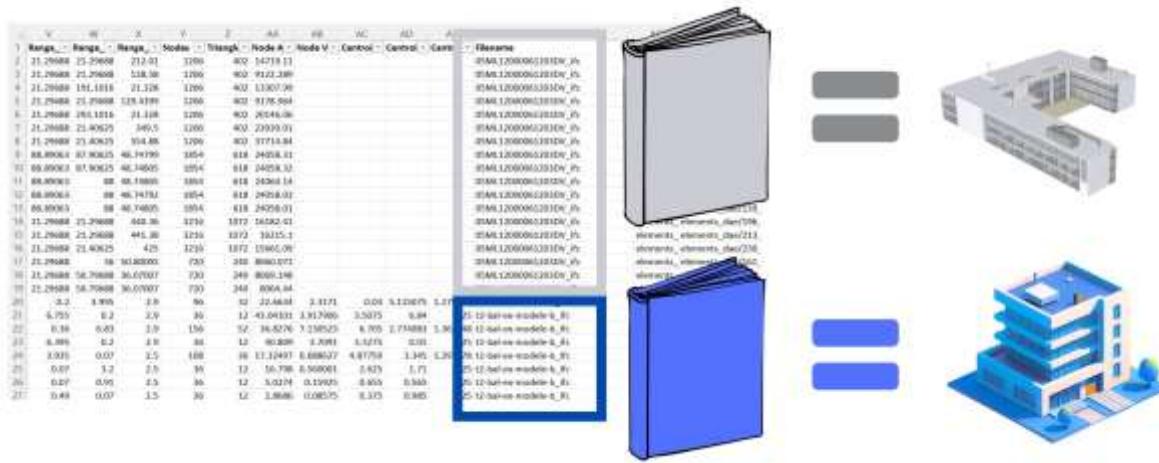
Lielie valodas modeļi (ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok) ir neironu tīkli, kas apmācīti, izmantojot milzīgus teksta datu apjomus no interneta, grāmatām, rakstiem un citiem avotiem. To galvenais uzdevums ir saprast cilvēka runas kontekstu un ģenerēt jēgpilnas atbildes.

Mūsdien LLM pamatā ir transformatora arhitektūra, ko 2017. gadā piedāvāja Google pētnieki [85]. Šīs arhitektūras galvenais komponents ir uzmanības mehānisms, kas ļauj modelim ļemt vērā attiecības starp vārdiem neatkarīgi no to atrašanās vietas tekstā.

LLM mācīšanās process ir ļoti līdzīgs tam, kā cilvēki mācās valodu, tikai miljoniem reižu lielāks. Modelis analizē miljardiem vārdu un izteicienu piemēru, identificējot valodas struktūras un semantisko pāreju loģikas modeļus. Viss teksts tiek sadalīts žetonos - minimālās semantiskās vienībās (vārdos vai to daļās), kas pēc tam tiek pārveidotas par vektoriem daudzdimensiju telpā ( 8.2-2. attēls). Šīs vektoru reprezentācijas ļauj mašīnai "saprast" slēptās attiecības starp jēdzieniem, nevis vienkārši operēt ar tekstu kā ar simbolu secību.

Lielie valodas modeļi nav tikai rīki teksta ģenerēšanai. Tie spēj atpazīt nozīmi, atrast saiknes starp jēdzieniem un strādāt ar datiem, pat ja tie ir iesniegti dažādos formātos. Galvenais, lai informācija tikuca sadalīta saprotamos modeļos un attēlotā kā žetoni, ar kuriem LLM var strādāt.

To pašu pieeju var piemērot arī būvniecības projektiem. Ja mēs domājam par projektu kā par sava veida tekstu, kurā katrā ēka, elements vai konstrukcija ir simbols, mēs varam sākt apstrādāt šādu informāciju līdzīgā veidā. Būvprojektus var salīdzināt ar grāmatām, kas ir sakārtotas kategorijās, nodaļās un rindkopu grupās, kuras sastāv no minimāliem žetoniem - būvprojekta elementiem ( 3.3-1. attēls). Tulkojot datu modeļus strukturētā formātā, strukturētus datus mēs varam pārvērst arī vektoru bāzēs ( 8.2-2. attēls), kas ir ideāls avots mašīnmācīšanās un tādām tehnoloģijām kā LLM.



Attēls 3.3-1 Būvprojekta elements ir kā zīmogs tekstā: minimālā vienība, no kuras tiek veidotas visa projekta sadaļu (kategoriju) grupas (rindkopas).

Ja būvprojekts ir digitalizēts un tā elementi ir attēloti kā žetoni vai vektori, ir iespējams tiem piekļūt dabiskā valodā, nevis izmantojot stingrus formālus pieprasījumus. Šeit izpaužas viena no galvenajām LLM priekšrocībām - spēja saprast vaicājuma nozīmi un sasaistīt to ar attiecīgajiem datiem.

Inženierim vairs nav jāraksta SQL vai Python kods, lai iegūtu vajadzīgos datus - viņš var vienkārši, izprotot LLM un datu struktūru, formulēt uzdevumu parastajā veidā: "Atrodiet visas dzelzsbetona konstrukcijas, kuru betona klase ir augstāka par B30, un aprēķiniet to kopējo tilpumu". Modelis atpazīs vaicājuma nozīmi, pārveidos to mašīnlasāmā formā, atradīs datus (sagrupēs un pārveidos) un atgriezīs galīgo rezultātu.

Dokumenti, tabulas, projektu modeļi tiek pārvērsti vektoru attēlojumā (iestrādāšana) un saglabāti datubāzē. Kad lietotājs uzdod jautājumu, arī vaicājums tiek pārvērts vektorā, un sistēma atrod atbilstošākos datus. Tas ļauj LLM paļauties ne tikai uz savām apmācītajām zināšanām, bet arī uz faktiskajiem uzņēmuma datiem, pat ja tie ir parādījušies jau pēc modeļa apmācības beigām.

Viena no svarīgākajām LLM priekšrocībām būvniecībā ir iespēja ģenerēt programmas kodu. Tā vietā, lai tehnisko uzdevumu nodotu programmētājam, speciālisti var uzdevumu aprakstīt dabiskajā valodā, un modelis izveidos nepieciešamo kodu, ko var izmantot (kopējot no tērzēšanas), veidojot procesu automatizācijas kodu. LLM -modeļi ļauj speciālistiem bez padziļinātām programmēšanas zināšanām sniegt ieguldījumu uzņēmuma biznesa procesu automatizācijā un uzlabošanā.



**Attēls 3.3-2 LLM nodrošina lietotājiem iespēju rakstīt kodu un iegūt rezultātus bez nepieciešamajām programmēšanas prasmēm.**

Saskaņā ar Wakefield Research veiktā un SAP sponsorētā 2024. gada pētījuma [36] datiem, kurā tika aptaujāti 300 augstākā līmeņa vadītāji ASV uzņēmumos, kuru gada ieņēmumi ir vismaz 1 miljards ASV dolāru: 52% augstākā līmeņa vadītāju uzticas mākslīgajam intelektam, lai analizētu datus un sniegtu ieteikumus lēmumu pieņemšanai. Vēl 48% izmanto AI, lai identificētu iepriekš neņemtus vērā riskus, un 47% izmanto AI, lai ieteiku alternatīvus plānus. Turklāt 40% izmanto AI jaunu produktu izstrādē, budžeta plānošanā un tirgus izpētē. Pētījums arī parādīja AI pozitīvo ietekmi uz personīgo dzīvi - 39% respondentu ziņoja par uzlabotu darba un privātās dzīves līdzsvaru, 38% - par uzlabotu garīgo veselību un 31% - par zemāku stresa līmeni.

Tomēr, neraugoties uz to spēku, LLM joprojām ir instruments, kuru ir svarīgi izmantot apzināti. Tāpat kā jebkurai tehnoloģijai, arī tiem ir ierobežojumi. Viena no pazīstamākajām problēmām ir tā sauktās "halucinācijas" - gadījumi, kad modelis pārliecinoši sniedz ticamu, bet faktiski nepareizu atbildi. Tāpēc ir ļoti svarīgi saprast, kā modelis darbojas: kādus datus un datu modeļus tas var interpretēt bez kļūdām, kā tas interpretē vaicājumus un no kurienes tas iegūst informāciju. Tāpat ir vērts atcerēties, ka LLM zināšanas ir ierobežotas ar tā apmācības datumu, un bez saiknes ar ārējiem datiem modelis var neņemt vērā pašreizējās normas, standartus, cenas vai tehnoloģijas.

Šo problēmu risinājums ir regulāri atjaunināt vektoru datubāzes, pieslēgties attiecīgajiem avotiem un izstrādāt autonomus mākslīgā intelekta aģentus, kas ne tikai atbild uz jautājumiem, bet proaktīvi izmanto datus apmācībai, pārvalda uzdevumus, identificē riskus, iesaka optimizācijas iespējas un uzrauga projekta izpildi.

Pāreja uz LLM - saskarnēm būvniecībā nav tikai tehnoloģisks jaunums. Tā ir paradigmas maiņa, kas likvidē šķēršļus starp cilvēkiem un datiem. Tā ir iespēja strādāt ar informāciju tikpat vienkārši, kā mēs runājam viens ar otru, un joprojām iegūt precīzus, pārbaudītus un izmantojamus rezultātus.

Tie uzņēmumi, kas šādus rīkus sāks izmantot agrāk nekā citi, iegūs ievērojamas konkurences priekšrocības. Tas ietver darba paātrināšanu, izmaksu samazināšanu un dizaina risinājumu kvalitātes

uzlabošanu, pateicoties ātrai piekļuvei datu analīzei un spējai ātri atrast atbildes uz sarežģītiem jautājumiem. Taču jāņem vērā arī drošības jautājumi. LLM mākoņpakaļpojumu izmantošana var būt saistīta ar datu noplūdes risku. Tāpēc organizācijas arvien biežāk meklē alternatīvus risinājumus, kas ļauj LLM rīkus izvietot pašu infrastruktūrā - lokāli, nodrošinot pilnīgu informācijas aizsardzību un kontroli.

## Vietējo LLM izmantošana sensitīviem uzņēmuma datiem

Pirmais tērzesanas-LLMs parādīšanās 2022. gadā iezīmēja jaunu posmu mākslīgā intelekta attīstībā. Tomēr uzreiz pēc šo modeļu plašas ieviešanas radās pamatots jautājums: cik droša ir ar uzņēmumu saistītu datu un pieprasījumu nodošana mākonī? Lielākā daļa mākoņa valodas modeļu uzkrāja saziņas vēsturi un augšupielādētos dokumentus savos serveros, un uzņēmumiem, kas strādā ar sensitīvu informāciju, tas bija nopietns šķērslis mākslīgā intelekta ieviešanai.

Viens no ilgtspējīgākajiem un loģiskākajiem šīs problēmas risinājumiem ir atklātā koda LLM izvietošana lokāli, uzņēmuma IT infrastruktūrā. Atšķirībā no mākoņpakaļpojumiem lokālie modeļi darbojas bez interneta savienojuma, nenodod datus uz ārējiem serveriem un nodrošina uzņēmumiem pilnīgu kontroli pār informāciju

*Labākais atvērtais modelis [Open Source LLM] pašlaik ir salīdzināms ar slēgtajiem modeļiem [piemēram, ChatGPT, Claude], taču ar aptuveni viena gada atpalicību [77].*

- Bens Kotjē, bezpeļņas pētniecības organizācijas Epoch AI vadošais pētnieks, 2024. gads.

Lielākie tehnoloģiju uzņēmumi ir sākuši padarīt savus LLM pieejamus vietējai lietošanai. Meta atvērtā koda LLaMA sērija un strauji augošais DeepSeek projekts no Ķīnas bija piemēri pārejai uz atvērto arhitektūru. Līdztekus tiem arī Mistral un Falcon ir izlaiduši jaudīgus modeļus, kas ir brīvi no patentētu platformu ierobežojumiem. Šīs iniciatīvas ir ne tikai paātrinājušas globālā mākslīgā intelekta attīstību, bet arī sniegušas uzņēmumiem, kas rūpējas par privātumu, reālas alternatīvas neatkarības, elastīguma un drošības atbilstības nodrošināšanai.

Uzņēmumu vidē, jo īpaši būvniecības nozarē, datu aizsardzība ir ne tikai ērtības jautājums, bet arī regulējuma ievērošanas jautājums. Darbs ar konkursa dokumentiem, tāmēm, rasējumiem un konfidenciālu saraksti prasa stingru kontroli. Un tieši šeit vietējais LLM sniedz nepieciešamo garantiju, ka dati paliek uzņēmuma perimetrā.

	Cloud LLMs (OpenAI, Claude)	Local LLMs (DeepSeek, LLaMA)
<b>Data Control</b>	Data is transmitted to third parties	Data remains within the company's network
<b>License</b>	Proprietary, paid	Open-source (Apache 2.0, MIT)
<b>Infrastructure</b>	Requires internet	Operates in an isolated environment
<b>Customization</b>	Limited	Full adaptation to company needs
<b>Cost</b>	Pay-per-token/request	One-time hardware investment + maintenance costs
<b>Scalability</b>	Easily scalable with cloud resources	Scaling requires additional local hardware
<b>Security &amp; Compliance</b>	Risk of data leaks, may not meet strict regulations (GDPR, HIPAA)	Full compliance with internal security policies
<b>Performance &amp; Latency</b>	Faster inference due to cloud infrastructure	Dependent on local hardware, may have higher latency
<b>Integration</b>	API-based integration, requires internet access	Can be tightly integrated with on-premise systems
<b>Updates &amp; Maintenance</b>	Automatically updated by provider	Requires manual updates and model retraining
<b>Energy Consumption</b>	Energy cost is covered by provider	High power consumption for inference and training
<b>Offline Availability</b>	Not available without an internet connection	Works completely offline
<b>Inference Cost</b>	Pay-per-use model (cost scales with usage)	Fixed cost after initial investment

**Attēls 3.3-3 Vietējie modeļi nodrošina pilnīgu kontroli un drošību, savukārt mākoņrisinājumi piedāvā vienkāršu integrāciju un automātiskus atjauninājumus.**

Vietējā atvērtā koda LLM galvenie ieguvumi:

- Pilnīga datu kontrole. Visa informācija paliek uzņēmumā, novēršot nesankcionētu piekļuvi un datu noplūdi.
- Bezsaistes darbība. Nav atkarības no interneta savienojuma, kas ir īpaši svarīgi izolētām IT infrastruktūrām. Tas nodrošina nepārtrauktu darbību arī sankciju vai bloķētu mākoņpakalpojumu gadījumā.
- Pielietojuma elastīgums. Šo modeli var izmantot teksta ģenerēšanai, datu analīzei, programmas koda rakstīšanai, projektēšanas atbalstam un biznesa procesu pārvaldībai.
- Pielāgošana uzņēmuma mērķiem. LLM var apmācīt par iekšējiem dokumentiem, kas ļauj nemt vērā uzņēmuma darba specifiku un nozares īpatnības. Vietējo LLM var savienot ar CRM,

ERP vai BI platformām, kas ļauj automatizēt klientu pieprasījumu analīzi, pārskatu izveidi vai pat tendenču prognozēšanu.

DeepSeek bezmaksas un atvērtā pirmkoda modeļa -R1-7B izvietošana serverī, lai tam varētu piekļūt visa lietotāju komanda, par 1000 ASV dolāriem mēnesī var potenciāli izmaksāt mazāk nekā gada maksa par mākoņa API, piemēram, ChatGPT vai Claude, un ļauj uzņēmumiem pilnībā kontrolēt savus datus, novērš to pārsūtīšanu uz internetu un palīdz izpildīt tādas regulatīvās prasības kā GDPR

Citās nozarēs vietējie LLMs jau maina savu pieeju automatizācijai. Atbalsta pakalpojumu jomā tie atbild uz biežiem klientu pieprasījumiem, samazinot operatoru darba slodzi. Personāla nodaļas tie analizē CV un atlasa atbilstošus kandidātus. E-komercijā tie ġenerē personalizētus piedāvājumus, neatklājot lietotāja datus.

Līdzīga ietekme gaidāma arī būvniecības nozarē. Integrējot LLM ar projekta datiem un standartiem, ir iespējams paātrināt dokumentācijas sagatavošanu, automatizēt tāmju sagatavošanu un prognozējamo izmaksu analīzi. LLM izmantošana kopā ar strukturētām tabulām un datu struktūrām klūst par īpaši daudzsološu jomu.

## Pilnīga kontrole pār AI uzņēmumā un kā izvietot savu LLM

Mūsdienīgi rīki ļauj uzņēmumiem lokāli izvietot lielu valodas modeli (LLM) tikai dažu stundu laikā. Tas nodrošina pilnīgu kontroli pār datiem un infrastruktūru, novēršot atkarību no ārējiem mākoņpakkalpojumiem un samazinot informācijas noplūdes risku. Šis risinājums ir īpaši piemērots organizācijām, kas strādā ar sensitīvu projektu dokumentāciju vai komerciāli sensitīviem datiem.

Atkarībā no uzdevumiem un resursiem ir pieejami dažādi izvietošanas scenāriji, sākot no gataviem risinājumiem līdz elastīgākām un mērogojamām arhitektūrām. Viens no vienkāršākajiem rīkiem ir Ollama, kas ļauj palaist valodas modeļus burtiski ar vienu klikšķi, bez nepieciešamības pēc dzīlām tehniskām zināšanām. Ātrs sākums ar Ollama:

1. Lejupielādējiet savai operētājsistēmai (Windows / Linux / macOS) paredzēto izplatījumu no oficiālās vietnes: [ollama.com](https://ollama.com).
2. Uzstādīt modeli, izmantojot komandrindi. Piemēram, *Mistral* modelim:

```
ollama palaist mistral
```

3. Pēc modeļa palaišanas tas ir gatavs darbam - varat sūtīt teksta pieprasījumus, izmantojot termināli, vai integrēt to citos rīkos. Palaidiet modeli un izpildiet vaicājumu:

```
ollama palaist mistral "Kā izveidot aprēķinu ar visiem resursiem darbam, lai uzstādītu 100 mm platu ģipškartona starpsienu?".
```

Tiem, kas dod priekšroku darbam pazīstamā vizuālajā vidē, ir pieejama LM Studio - bezmaksas lietojumprogramma ar saskarni, kas atgādina ChatGPT

- Instalējet LM Studio, lejupielādējot izplatīšanas komplektu no oficiālās vietnes - lmstudio.ai.
- Izmantojot iebūvēto katalogu, izvēlieties modeli (piemēram, Falcon vai GPT-Neo-X) un lejupielādējiet to.
- Darbs ar modeli, izmantojot intuitīvu saskarni, kas atgādina ChatGPT, bet ir pilnībā lokalizēta.

	Developer	Parameters	GPU Requirements (GB)	Features	Best For
Mistral 7B	Mistral AI	7	8 (FP16)	Fast, supports multimodal tasks (text + images), fully open-source code	Lightweight tasks, mobile devices, laptops
LLaMA 2	Meta	7–70	16–48 (FP16)	High text generation accuracy, adaptable for technical tasks, CC-BY-SA license	Complex analytical and technical tasks
Baichuan 7B/13B	Baichuan Intelligence	7–13	8–16 (FP16)	Fast and efficient, great for large data processing, fully open-source code	Data processing, automating routine tasks
Falcon 7B/40B	Technology Innovation Institute (TII)	7–40	8–32 (FP16)	Open-source, high performance, optimized for fast work	Workloads with limited computational resources
DeepSeek-V3	DeepSeek	671	1543 (FP16) / 386 (4-bit)	Multilingual, 128K token context window, balanced speed and accuracy	Large enterprises, SaaS platforms, multitasking scenarios
DeepSeek-R1-7B	DeepSeek	7	18 (FP16) / 4.5 (4-bit)	Retains 92% of R1 capabilities in MATH-500, local deployment support	Budget solutions, IoT devices, edge computing

Attēls 3.3-4 Populāru vietējo LLM modeļu salīdzinājums.

Modeļa izvēle ir atkarīga no prasībām attiecībā uz ātrumu, precizitāti un pieejamajām aparatūras iespējām ( 3.3-4. attēls). Mazie modeļi, piemēram, Mistral 7B un Baichuan 7B, ir piemēroti viegliem uzdevumiem un mobilajām ierīcēm, savukārt jaudīgiem modeļiem, piemēram, DeepSeek -V3, ir nepieciešami ievērojami skaitļošanas resursi, bet tie nodrošina augstu veikspēju un vairāku valodu atbalstu. Turpmākajos gados LLM tirgus strauji aug - mēs redzēsim arvien vairāk vieglo un specializēto modeļu. Tā vietā, lai vispārējas nozīmes LLM aptvertu visu cilvēku saturu, parādīsies modeļi, kas apmācīti, pamatojoties uz šauru jomu zināšanām. Piemēram, varam sagaidīt, ka parādīsies modeļi, kas paredzēti tikai inženiertehnisko aprēķinu, būvniecības tāmu vai datu apstrādei CAD formātos. Šādi specializēti modeļi būs ātrāki, precīzāki un drošāki lietošanā - īpaši profesionālajā vidē, kur svarīga ir augsta uzticamība un padzīlināta tematiskā informācija.

Kad vietējais LLM ir izveidots, to var pielāgot uzņēmuma specifiskajiem uzdevumiem. Šim nolūkam tiek izmantota precīzēšanas metode, ar kuras palīdzību modelis tiek papildus apmācīts, izmantojot iekšējos dokumentus, tehniskās instrukcijas, līgumu veidnes vai projekta dokumentāciju.

## RAG: Intelligent LLM - asistenti ar piekļuvi uzņēmuma datiem.

Nākamais posms LLM pielietojuma attīstībā uzņēmējdarbībā ir modeļu integrācija ar reālā laika uzņēmuma datiem. Šo pieeju sauc par RAG (Retrieval-Augmented Generation) - Retrieval-Augmented Generation. Šajā arhitektūrā valodas modelis kļūst ne tikai par dialoga saskarni, bet gan par pilnvērtīgu inteliģentu pašigu, kas spēj orientēties dokumentos, rasējumos, datubāzēs un sniegt precīzas, kontekstuālas atbildes.

Galvenā RAG priekšrocība ir iespēja izmantot uzņēmuma iekšējos datus bez nepieciešamības iepriekš apmācīt modeli, vienlaikus saglabājot augstu informācijas apstrādes precīzitāti un elastību.

RAG tehnoloģija apvieno divas galvenās sastāvdaļas:

- Izguve: modelis savienojas ar datu krātuvēm - dokumentiem, tabulām, PDF - failiem, rasējumiem - un iegūst attiecīgo informāciju pēc lietotāja pieprasījuma.
- Paplašināta ģenerēšana: pamatojoties uz iegūtajiem datiem, modelis ģenerē precīzu, pamatotu atbildi, ņemot vērā kontekstu un vaicājuma specifiku.

Lai palaistu LLM ar RAG atbalstu, ir jāveic daži soļi:

- **Datu sagatavošana:** apkopot nepieciešamos dokumentus, rasējumus, specifikācijas, tabulas. Tie var būt dažādos formātos un struktūrās, sākot no PDF līdz Excel.
- **Indeksēšana un vektorizācija:** izmantojot tādus rīkus kā LlamalIndex vai LangChain, dati tiek pārvērsti vektoru attēlojumos, kas ļauj atrast semantiskās saites starp teksta fragmentiem (vairāk par vektoru datubāzēm un lielu masīvu pārvēršanu vektoru attēlojumā, tostarp CAD projektiem, 8. daļā).
- Kad dati ir augšupielādēti, varat uzdot modelim jautājumus, un tas meklēs atbildes uzņēmuma sistēmā, nevis vispārējās zināšanās, kas apkopotas internetā.

Pieņemsim, ka uzņēmumam ir mape constructionsite\_docs, kurā tiek glabāti līgumi, instrukcijas, tāmes un tabulas. Izmantojot Python skriptu (3.3-5. attēls), mēs varam skenēt šo mapi un izveidot vektoru indeksāciju: katrs dokuments tiks pārveidots par vektoru kopu, kas atspoguļo teksta semantisko saturu. Tādējādi dokumenti kļūst par sava veida "nozīmju karti", pēc kuras modelis var efektīvi orientēties un atrast sakarības starp terminiem un frāzēm.

Piemēram, paraugs "atceras", ka vārdi "atgriešana" un "sūdzība" bieži vien ir atrodami līguma sadaļā, kas attiecas uz materiālu nosūtīšanu uz būvlaukumu. Tad, ja tiek uzdots jautājums, piemēram, "Kāds ir mūsu atgriešanas termiņš?". (3.3-5. attēls - koda 11. rindiņa) - LLM analizēs iekšējos dokumentus

un atradīs precīzu informāciju, darbojoties kā inteliģents palīgs, kas spēj izlasīt un saprast visu uzņēmuma failu saturu.



```

1 from llama_index import SimpleDirectoryReader, VectorStoreIndex
2
3 # Load documents from the folder
4 documents = SimpleDirectoryReader("constructionsite_docs").load_data()
5
6 # Creating a vector index for semantic search
7 index = VectorStoreIndex.from_documents(documents)
8
9 # Integration with LLM (e.g. Llama 3)
10 query_engine = index.as_query_engine()
11 response = query_engine.query("What are the return terms in the contracts?")
12 print(response)

```

### Attēls 3.3-5 LM nolasa mapi ar failiem - līdzīgi kā cilvēks to atver un meklē vajadzīgo dokumentu

Šo kodu var palaist jebkurā datorā, kurā ir instalēts Python. Par Python un IDE lietojumu, lai palaistu kodu, sīkāk runāsim nākamajā nodoļā.

Vietējā LLM izvietošana ir ne tikai tendence, bet stratēģisks risinājums uzņēmumiem, kas augstu vērtē drošību un elastīgumu. Tomēr LLM izvietošana, vai nu vietējos uzņēmuma datoros, vai izmantojot tiešsaistes risinājumus, ir tikai pirmsais solis. Lai izmantotu LLM iespējas reālu uzdevumu veikšanai, uzņēmumiem ir jāizmanto rīki, kas ļauj ne tikai saņemt tērzēšanas atbildes, bet arī saglabāt izveidoto loģiku koda veidā, ko var palaist ārpus LLM izmantošanas konteksta. Tas ir svarīgi risinājumu mērogošanai - pareizi organizēti procesi ļauj mākslīgā intelekta izstrādnes piemērot vairākiem projektiem vienlaikus vai pat visā uzņēmumā.

Šajā kontekstā liela nozīme ir piemērotas izstrādes vides (IDE) izvēlei. Mūsdienu programmēšanas rīki ļauj ne tikai izstrādāt uz LLM balstītus risinājumus, bet arī integrēt tos esošajos biznesa procesos, pārvēršot tos par automatizētiem ETL -Pipeline



## NODAĻA 3.4.

### IDE AR LLM ATBALSTU UN TURPMĀKĀS PROGRAMMĒŠANAS IZMAINAS

#### IDE izvēle: no LLM eksperimentiem līdz biznesa risinājumiem

Iesaistoties automatizācijas, datu analīzes un mākslīgā intelekta pasaulē, jo īpaši strādājot ar lieliem valodas modeļiem (LLM), ir ļoti svarīgi izvēlēties pareizo integrēto izstrādes vidi (IDE). Šī IDE būs jūsu galvenais darba rīks - vieta, kur tiks palaists LLM ģenerētais kods gan lokālajā datorā, gan uzņēmuma tīklā. IDE izvēle nosaka ne tikai jūsu darba ērtumu, bet arī to, cik ātri jūs varēsiet pāriet no eksperimentāliem LLM pieprasījumiem uz pilnvērtīgiem risinājumiem, kas iestrādāti reālos uzņēmējdarbības procesos.

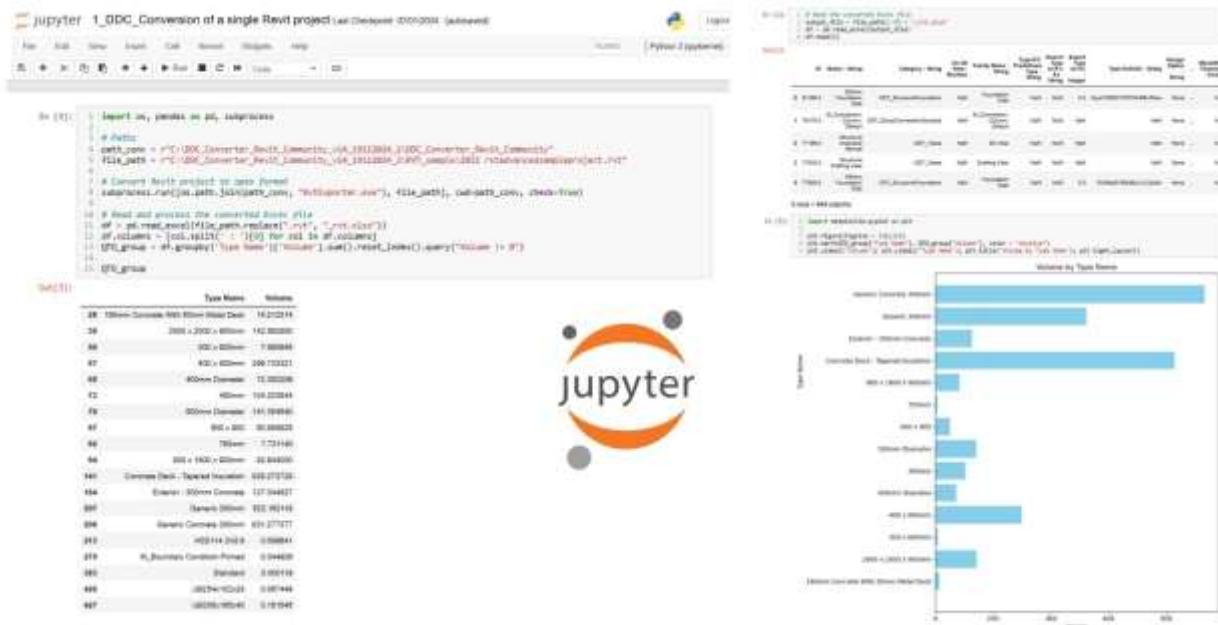
**IDE** (integrētā izstrādes vide) ir daudzpusīgs jūsu datora veidošanas elements procesu automatizācijai un datu apstrādei. Tā vietā, lai atsevišķi turētu zāģi, āmuru, urbjmašīnu un citus instrumentus, jums ir viena ierīce, kas var darīt visu - griezt, stiprināt, urbt un pat pārbaudīt materiālu kvalitāti. IDE programmētājiem ir vienota telpa, kurā var rakstīt kodu (pēc analogijas ar būvniecību - izveidot rasējumus), pārbaudīt tā darbu (būvmodeļa montāža), atrast klūdas (piemēram, pārbaudot konstrukciju stiprību būvniecībā) un palaist pabeigtu projektu (mājas nodošana ekspluatācijā).

Pārskats par populārākajām IDE:

- **PyCharm®** (JetBrains) ir jaudīga profesionāla IDE Python. Tā ir labi piemērota nopietniem projektiem, pateicoties lielam iebūvēto funkciju skaitam. Tomēr interaktīvo Jupyter failu (IPYNB) pamata atbalsts ir pieejams tikai maksas versijā, un iesācējiem saskarne var šķist pārlieku sarežģīta.

Faili ar paplašinājumu IPYNB (Interactive Python Notebook) ir interaktīvo Jupyter® piezīmjdatoru formāts (3.4-1. attēls), kurā kods, vizualizācijas un paskaidrojumi ir apvienoti vienā dokumentā. Šis formāts ir ideāli piemērots pārskatu, analītikas un mācību scenāriju veidošanai.

- **VS Code®** (Microsoft) ir ātrs, elastīgs un pielāgojams rīks ar bezmaksas IPYNB atbalstu un daudziem spraudņiem. Piemērots gan iesācējiem, gan profesionāļiem. Ľauj integrēt GitHub Copilot un valodas modeļu spraudņus, padarot to par lielisku izvēli mākslīgā intelekta un datu zinātnes projektiem.
- **Jupyter Notebook** - klasiska un populāra izvēle eksperimentiem un mācībām. Tā ļauj rakstīt kodu, pievienot paskaidrojumus un vizualizēt rezultātus vienotā saskarnē (3.4-1. attēls). Ideāli piemērots ātrai hipotēžu pārbaudei, darbam ar LLM un reproducējamu savvaļas datu analīzes soļu izveidei. Lai pārvaldītu atkarības un bibliotēkas, iesakām izmantot Anaconda Navigator, vizuālo saskarni Python vides pārvaldībai.



Attēls 3.4-1 Jupyter Notebook viens no ērtākajiem un populārākajiem rīkiem caurulvadu procesu izveidei.

- **Google Collab™** (un Kaggle platforma (9.2-5. attēls)) ir Jupyter mākoņbāzēta alternatīva, kas nodrošina bezmaksas piekļuvi GPU/TPU. Tas ir lielisks risinājums, lai sāktu darbu - nav nepieciešama lokālas programmatūras instalēšana un ir iespēja strādāt tieši no pārlūkprogrammas. Tā atbalsta integrāciju ar Google Drive un nesen arī ar Gemini (Google LLM).

	PyCharm	VS Code	Jupyter Notebook	Google Colab
Complexity	High	Medium	Low	Low
.ipynb support	Paid	Free	Built-in	Built-in
Copilots	Yes	Yes	Yes	Yes
Computing resources	Local	Local	Local	Cloud
For whom	Professionals	Universal	Beginners	Experimenters

**Attēls 3.4-2 IDE salīdzinājums: Jupyter Notebook viens no ērtākajiem un vienkāršākajiem rīkiem, lai izveidotu Pipeline procesus.**

IDE izvēle ir atkarīga no jūsu uzdevumiem. Ja vēlaties ātri sākt darbu ar mākslīgo intelektu, izmēģiniet Jupyter Notebook vai Google Collab. Nopietniem projektiem labāk izmantot PyCharm vai VS Code.

Galvenais ir sākt darbu. Mūsdienīgi rīki ļauj ātri pārvērst eksperimentus darbojošos risinājumos.

Visas iepriekš aprakstītās IDE ļauj izveidot datu apstrādes cauruļvadus, t. i., koda bloku moduļu ķēdes (ko varētu generēt LLM), no kuriem katrs ir atbildīgs, piemēram, par atsevišķu posmu:

- analītiskie scenāriji,
- informācijas ieguves ķēdes no dokumentiem,
- automātiskas atbildes, pamatojoties uz RAG,
- pārskatu un vizualizāciju ģenerēšana.

Pateicoties modulārajai struktūrai, katru posmu var attēlot kā atsevišķu bloku: datu ielāde→filtrēšana→ analīze→ vizualizēšana→ rezultātu eksportēšana. Šos blokus var izmantot atkārtoti, - pielāgot un salikt jaunās ķēdēs kā konstruktoru, tikai datiem.

Inženieriem, vadītājiem un analītiķiem tas paver iespēju dokumentēt lēmumu pieņemšanas loģiku koda veidā, ko var generēt ar LLM. Šāda pieeja palīdz paātrināt rutīnas uzdevumu izpildi, automatizēt tipiskas darbības un izveidot atkārtojamus procesus, kuros katrs solis ir skaidri dokumentēts un pārredzams visiem komandas locekļiem.

Automatizētie ETL Pipelines ( 7.2-3. attēls), Apache Airflow ( 7.4-4. attēls), Apache NiFi ( 7.4-5. attēls) un n8n ( 7.4-6. attēls) rīki procesu automatizācijas loģikas bloku veidošanai tiks sīkāk apskatīti grāmatas 7. un 8. daļā.

## **IDE ar LLM atbalstu un turpmākās programmēšanas izmaiņas**

Mākslīgā intelekta integrēšana izstrādes procesos maina programmēšanas ainavu. Mūsdienu vides vairs nav tikai teksta redaktori ar sintakses izcelšanu - tās kļūst par inteliģentiem paīgiem, kas spēj izprast projekta loģiku, papildināt kodu un pat izskaidrot, kā darbojas konkrēts koda fragments. Tirgū parādās produkti, kas izmanto mākslīgo intelektu, lai paplašinātu tradicionālās izstrādes robežas:

- **GitHub Copilot** (integrējas ar VS Code, PyCharm): mākslīgais intelekts - palīgs, kas generē kodu, pamatojoties uz komentāriem vai daļējiem aprakstiem, pārvēršot teksta mājienus pabeigtos risinājumos.
- **Kursors** (VS Code atvasinājums ar AI -kernel): ļauj ne tikai pabeigt rakstīt kodu, bet arī uzdot jautājumus projektam, meklēt atkarības un mācīties no koda bāzes.
- **JetBrains AI Assistant**: JetBrains IDE (tostarp PyCharm) spraudnis, kura funkcija ir izskaidrot sarežģītu kodu, optimizēt un izveidot testus.
- **Amazon CodeWhisperer**: Copilot analogs, kas koncentrējas uz drošību un Amazon AWS pakalpojumu atbalstu.

Turpmākajos gados programmēšana piedzīvos būtiskas pārmaiņas. Galvenā uzmanība tiks pievērsta nevis rutīnas koda rakstīšanai, bet gan modeļu un datu arhitektūras projektēšanai - izstrādātāji vairāk iesaistīsies sistēmas projektēšanā, savukārt mākslīgais intelekts pārņems šablonu uzdevumus: koda ģenerēšanu, testus, dokumentāciju un pamatfunkcijas. Programmēšanas nākotne ir cilvēku un mākslīgā intelekta sadarbība, kurā mašīnas pārņems tehnisko rutīnu, bet cilvēki koncentrēsies uz

radošumu.

Dabiskās valodas programmēšana kļūs par ikdienišķu parādību. IDE personalizācija sasniegs jaunu līmeni - izstrādes vides mācīties pielāgoties lietotāja darba stilam un uzņēmumiem, paredzot modeļus, piedāvājot kontekstuālus risinājumus un mācoties no iepriekšējiem projektiem

Tas neatcel izstrādātāja lomu, taču radikāli pārveidos to: no koda rakstīšanas uz zināšanu, kvalitātes un procesu pārvaldību. Šī evolūcija ietekmēs arī biznesa izlūkošanu, kur pārskatu, vizualizāciju un lēmumu atbalsta lietojumprogrammu izveide arvien vairāk tiks veikta, ģenerējot kodu un loģiku ar mākslīgā intelekta un LLM, tērzēšanas un aģentu saskarņu paīdzību.

Kad uzņēmums ir izveidojis LLM tērzēšanu un izvēlējies piemērotu izstrādes vidi, nākamais svarīgais solis ir organizēt datus. Šis process ietver informācijas iegūšanu no dažādiem avotiem, tās attīrīšanu, pārveidošanu strukturētā formā un integrēšanu uzņēmuma sistēmās.

Mūsdienu datu centriskās piejas datu pārvaldībai galvenais mērķis ir datus apkopot vienotā universālā formā, kas ir saderīga ar daudziem rīkiem un lietojumprogrammām. Strukturēšanas procesu un strukturētu datu apstrādei ir nepieciešamas specializētas bibliotēkas. Viena no jaudīgākajām, elastīgākajām un populārākajām ir Pandas bibliotēka Python. Tā ļauj ērti apstrādāt tabulāros datus: filtrēt, grupēt, attīrīt, pievienot, apkopot un veidot pārskatus.

## **Python Pandas: neaizstājams rīks darbam ar datiem**

Pandas ieņem īpašu vietu datu analīzes un automatizācijas pasaulē. Tā ir viena no populārākajām un visplašāk izmantotajām Python programmēšanas valodas bibliotēkām [86], kas paredzēta darbam ar strukturētiem datiem.

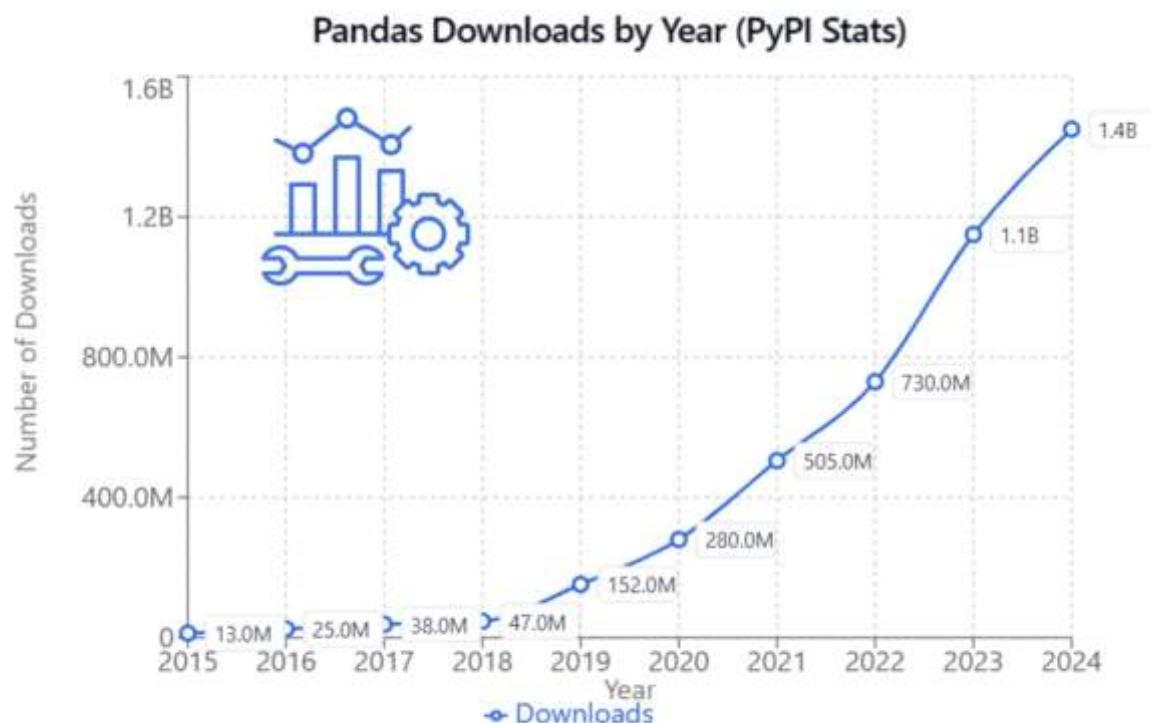
Bibliotēka ir kā gatavu rīku kopums: funkcijas, moduļi, klases. Tāpat kā būvlaukumā katru reizi nav jāizgudro āmurs vai līmeņrādis, tā arī programmēšanas jomā bibliotēkas ļauj ātri atrisināt problēmas, neizgudrojot no jauna pamattfunkcijas un risinājumus.

Pandas ir atvērtā pirmkoda Python bibliotēka , kas nodrošina augstas veiktspējas un intuitīvas datu struktūras, jo īpaši DataFrame, kas ir universāls formāts darbam ar tabulām. Pandas ir Šveices nazis analītiķiem, inženieriem un izstrādātājiem, kas strādā ar datiem.

Python ir augsta līmeņa programmēšanas valoda ar vienkāršu sintaksi, ko aktīvi izmanto analīzē, automatizācijā, mašīnmācībā un tīmekļa izstrādē. Tās popularitāti nosaka tās koda lasāmība, daudzplatformu raksturs un bagātā bibliotēku ekosistēma. Līdz šim Python ir izveidotas vairāk nekā 137 000 atvērtā pirmkoda pakotņu [87], un šis skaits gandrīz katru dienu turpina pieaugt. Katra šāda bibliotēka ir sava veida gatavu funkciju krātuve: no vienkāršām matemātiskām operācijām līdz sarežģītiem rīkiem attēlu apstrādei, lielo datu analīzei, neironu tīkliem un integrācijai ar ārējiem pakalpojumiem.

Citiem vārdiem sakot, iedomājieties, ka jums ir brīva un atklāta piekļuve simtiem tūkstošu gatavu programmatūras risinājumu - bibliotēku un rīku, kurus varat tieši integrēt savos uzņēmējdarbības procesos. Tas ir kā milzīgs automatizācijas, analīzes, vizualizācijas, integrācijas un citu lietojumprogrammu katalogs, kas ir pieejams uzreiz pēc Python instalēšanas.

Pandas ir viena no populārākajām Python ekosistēmas pakotnēm. Pandas bibliotēkas vidējais lejupielādes skaits 2022. gadā sasniedza 4 miljonus lejupielādes dienā (3.4-3. attēls), savukārt 2025. gada sākumā šis rādītājs pieauga līdz 12 milioniem lejupielādes dienā, atspoguļojot tās pieaugošo popularitāti un plašo izmantošanu datu analīzē un LLM čatos [86].

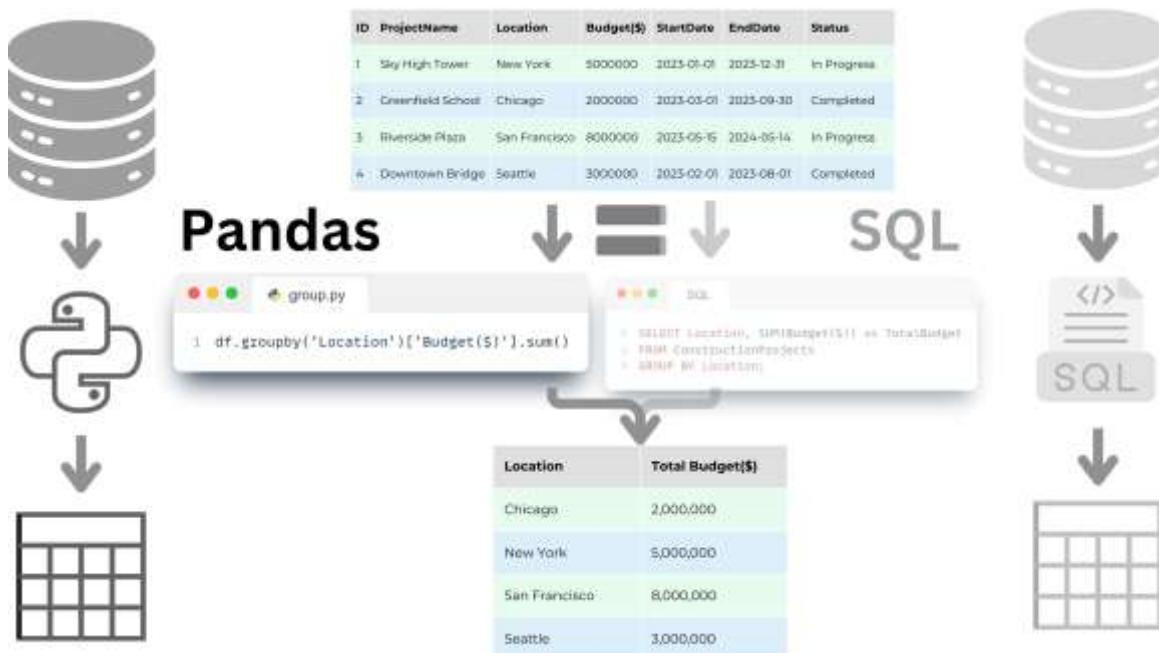


Attēls 3.4-3 Pandas ir viena no visvairāk lejupielādētajām bibliotēkām. Tās gada lejupielādes 2024. gadā pārsniedza 1,4 miljardus.

Pandas bibliotēkas vaicājumu valoda ir līdzīga SQL vaicājumu valodai, par kuru mēs runājām nodaļā "Relāciju datu bāzes un SQL vaicājumu valoda".

Analītikas un strukturētu datu pārvaldības pasaulē Pandas izceļas ar savu vienkāršību, ātrumu un jaudu, nodrošinot lietotājiem plašu rīku klāstu, lai efektīvi analizētu un apstrādātu informāciju.

Abi rīki - SQL un Pandas - nodrošina jaudīgas datu apstrādes iespējas, īpaši salīdzinājumā ar tradicionālo Excel. Tie atbalsta tādas operācijas kā atlase, filtrēšana (3.4-4. attēls), vienīgā atšķirība ir tā, ka SQL ir optimizēts darbam ar relāciju datu bāzēm, bet Pandas apstrādā datus operatīvajā atmiņā, kas ļauj to darbināt jebkurā datorā, bez nepieciešamības veidot datu bāzes un izvietot atsevišķu infrastruktūru.

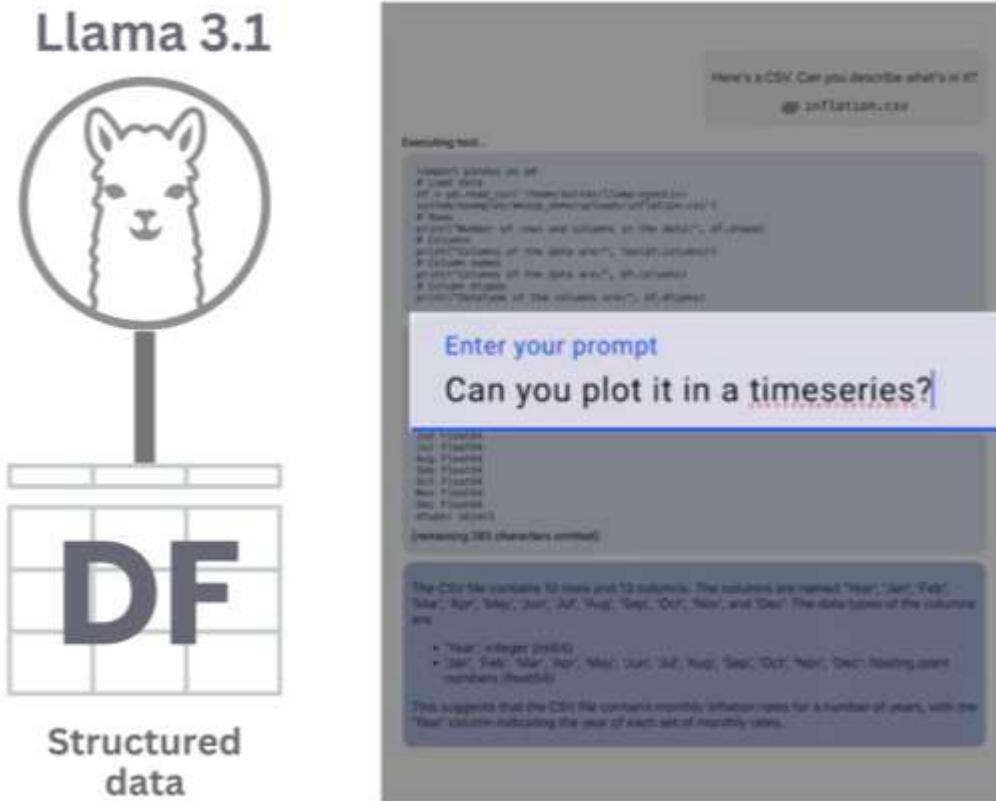


Attēls 3.4-4 Pandas, atšķirībā no SQL, ir elastīgs, lai strādātu ar dažādiem datu formātiem, ne tikai datubāzēm.

Pandas bieži vien tiek dota priekšroka zinātniskajiem pētījumiem, procesu automatizācijai, cauruļvadu izveidei (tostarp ETL) un datu apstrādei Python, savukārt SQL ir datubāzes pārvaldības standarts, un to bieži izmanto uzņēmumu vidēs, lai apstrādātu lielus datu apjomus.

Python programmēšanas valodas Pandas bibliotēka ļauj veikt ne tikai pamatdarbības, piemēram, tabulu lasīšanu un rakstīšanu, bet arī sarežģītākus uzdevumus, tostarp datu apvienošanu, datu grupēšanu un sarežģītus analītiskus aprēķinus.

Mūsdienās Pandas bibliotēka tiek izmantota ne tikai akadēmiskajos pētījumos un biznesa analīzē, bet arī kopā ar LLM -modeļiem. Piemēram, Meta® nodaļa (Facebook™), 2024. gadā publicējot jauno atvērtā pirkoda modeli LLaMa 3.1, īpašu uzmanību pievērsa darbam ar strukturētiem datiem, par vienu no galvenajiem un pirmajiem gadījumiem tā izlaidumā padarot tieši strukturētu datu rāmju apstrādi (3.4-5. attēls) CSV formātā un integrāciju ar Pandas bibliotēku tieši tērzēšanas laikā.



Attēls 3.4-5 Viens no Meta komandas pirmajiem un galvenajiem gadījumiem, kas tika prezentēts LlaMa 3.1 2024. gadā, bija lietojumprogrammu veidošana, izmantojot Pandas.

*Pandas ir būtisks rīks miljoniem datu zinātnieku, kas apstrādā un sagatavo datus ģeneratīvajam mākslīgajam intelektam. Pandas paātrināšana, nemainot kodu, būs milzīgs solis uz priekšu. Datu zinātnieki varēs apstrādāt datus dažu minūšu, nevis stundu laikā un iegūt par kārtām vairāk datu ģeneratīvā mākslīgā intelekta modeļu apmācībai [88].*

- Jensen Huang, NVIDIA dibinātājs un izpilddirektors.

Izmantojot Pandas, ir iespējams pārvaldīt un analizēt datu kopas, kas ievērojami pārsniedz Excel iespējas. Kamēr Excel parasti spēj apstrādāt līdz 1 miljonam datu rindu, Pandas var viegli apstrādāt datu kopas (9.1-2. attēls, 9.1-10. attēls), kas satur desmitiem miljonu rindu [89]. Šī spēja ļauj lietotājiem veikt sarežģītu datu analīzi un lielu datu kopu vizualizāciju, nodrošinot padziļinātu ieskatu un atvieglojot uz datiem balstītu lēmumu pieņemšanu. Turklāt Pandas ir spēcīgs kopienas atbalsts [90]: simtiem miljonu izstrādātāju un analītiķu visā pasaule (Kaggle.com, Google Collab, Microsoft® Azure™ Notebooks, Amazon SageMaker) to izmanto tiešsaistē vai bezsaistē katru dienu, nodrošinot lielu skaitu pieejamu risinājumu jebkurai uzņēmējdarbības problēmai.

Lielākās daļas Python analītisko procesu pamatā ir strukturēta datu forma, ko sauc par DataFrame, ko

nodrošina Pandas bibliotēka. Tas ir jaudīgs un elastīgs rīks tabulāro datu organizēšanai, analīzei un vizualizēšanai.

## DataFrame: universāls tabulāro datu formāts

DataFrame ir centrālā Pandas bibliotēkas struktūra, kas ir divdimensiju tabula ( 3.4-6. attēls), kurā rindas atbilst atsevišķiem objektiem vai ierakstiem, bet kolonas atbilst to īpašībām, parametriem vai kategorijām. Šī struktūra vizuāli atgādina Excel izklājlapas, bet ir daudz labāka elastības, mērogojamības un funkcionalitātes ziņā.

**DataFrame** ir veids, kā attēlot un apstrādāt tabulāros datus, kas glabājas datora operatīvajā atmiņā.

DataFrame ir veids, kā attēlot un apstrādāt tabulāros datus, kas glabājas datora RAM. Tabulā rindas var atspoguļot, piemēram, būvprojekta elementus, bet kolonas - to īpašības: kategorijas, izmērus, koordinātas, izmaksas, nosacījumus utt. Turklāt šāda tabula var saturēt gan informāciju par vienu projektu ( 4.1-13. attēls), gan datus par miljoniem objektu no tūkstošiem dažādu projektu ( 9.1-10. attēls). Pateicoties vektorizētajām Pandas operācijām, ir viegli un ātri filtrēt, grupēt un apkopot šādus informācijas apjomus.

The diagram illustrates a DataFrame as a structured data table for a construction project. The table has the following structure:

ID	Name	Category	Family Name	Height	BoundingBoxMin_X	BoundingBoxMin_Y	BoundingBoxMin_Z	Level
431144	Single-Flush	OST_Doors	Single-Flush	6.88976378	20.1503	-10.438	9.84252	Level 1
431198	Single-Flush	OST_Doors		6.88976378	13.2281	-1.1207	9.84252	Level 2
457479	Single Window	OST_Windows	Single Window	8.858267717	-11.434	-11.985	9.80971	Level 2
485432	Single Window	OST_Windows	Single Window	8.858267717	-11.434	4.25986	9.80971	Level 2
490150	Single-Flush	OST_Doors	Single-Flush	6.88976378	-1.5748	-2.9565	-1E-16	Level 1
493697	Basic Wall	OST_Walls	Basic Wall		-38.15	20.1656	-4.9213	Level 1
497540	Basic Wall	OST_Walls	Basic Wall		-4.5212	0.0708	9.84252	Level 1

Attēls 3.4-6 Construction Project as a DataFrame ir divdimensiju tabula ar elementiem rindās un atribūtiem kolonnās.

Nvidia lēš, ka jau šobrīd līdz pat 30 procentiem no visiem skaitļošanas resursiem tiek izmantoti strukturētu datu - datu kadru - apstrādei, un šī daļa turpina pieaugt.

Katrā uzņēmumā, iespējams, tiek apstrādāta trešdaļa no pasaules skaitļošanas darbiem. Datu apstrāde un dati lielākajā daļā uzņēmumu tiek apstrādāti DataFrame, tabulas formātā

- Jensen Huang, Nvidia izpilddirektors [91]

Uzskaitīsim dažas galvenās DataFrame funkcijas Pandas:

- **Kolonnas:** datu rāmī dati ir sakārtoti kolonnās, no kurām katrai ir unikāls nosaukums. Atribūtu kolonnas var saturēt dažādu tipu datus, līdzīgi kā datu bāzu kolonas vai tabulas kolonnas.
- **Pandas Sērija** ir viendimensiju datu struktūra programmā Pandas, kas līdzinās sarakstam vai slejām tabulā, kur katrai vērtībai atbilst cits indekss.
- Pandas sērijā ir vairāk nekā 400 atribūtu un metožu, kas padara darbu ar datiem neticami elastīgu. Jūs varat tieši piemērot kolonnai vienu no četriem simtiem pieejamo funkciju, veikt matemātiskas darbības, filtrēt datus, aizstāt vērtības, strādāt ar datumiem, virknēm un daudz ko citu. Turklāt Series atbalsta vektorizētas operācijas, kas ievērojami paātrina lielu datu kopu apstrādi salīdzinājumā ar cikliskajiem aprēķiniem. Piemēram, varat viegli reizināt visas vērtības ar skaitli, aizstāt trūkstošos datus vai piemērot sarežģītas transformācijas, nerakstot sarežģītus ciklus.
- **Rindas:** datu rāmī var indeksēt ar unikālām vērtībām. Šis indekss ļauj ātri mainīt un pielāgot datus konkrētās rindās.
- **Indekss:** Pēc noklusējuma, kad izveidojat DataFrame Pandas, katrai rindai tiek piešķirts indekss no 0 līdz N-1 (kur N ir visu DataFrame rindu skaits). Tomēr indeksu var mainīt, lai iekļautu īpašus apzīmējumus, piemēram, datumus vai unikālas īpašības.
- Datu rāmja rindu **indeksēšana** nozīmē, ka katrai rindai tiek piešķirts unikāls nosaukums vai etikete, ko sauc par datu rāmja indeksu.
- **Datu tipi:** DataFrame atbalsta dažādus datu tipus, tostarp: `int`, `float`, `bool`, `datetime64` un `object` teksta datiem. Katrai DataFrame kolonnai ir sava datu tips, kas nosaka, kādas darbības var veikt ar tās saturu.
- **Datu apstrādes operācijas:** DataFrame atbalsta plašu datu apstrādes operāciju klāstu, tostarp apkopošanu (`groupby`), apvienošanu (`merge` un `join`), konkatenāciju (`concat`), sadalīšanu-aplikāciju-kombināciju un daudzas citas datu pārveidošanas metodes.
- **Izmēra manipulācijas:** DataFrame ļauj pievienot un noņemt kolonas un rindas, padarot to par dinamisku struktūru, ko var mainīt atbilstoši datu analīzes vajadzībām.
- **datu vizualizēšana:** izmantojot iebūvētās vizualizācijas metodes vai saskarni ar populārām datu vizualizācijas bibliotēkām, piemēram, Matplotlib vai Seaborn, DataFrame var viegli pārvērst grafikos un diagrammās, lai datus attēlotu grafiski.
- **Datu ievade un izvade:** Pandas nodrošina funkcijas datu lasīšanai, importēšanai un eksportēšanai dažādos failu formātos, piemēram, CSV, Excel, JSON, HTML un SQL, tādējādi DataFrame var kļūt par centrālo datu vākšanas un izplatīšanas centru.

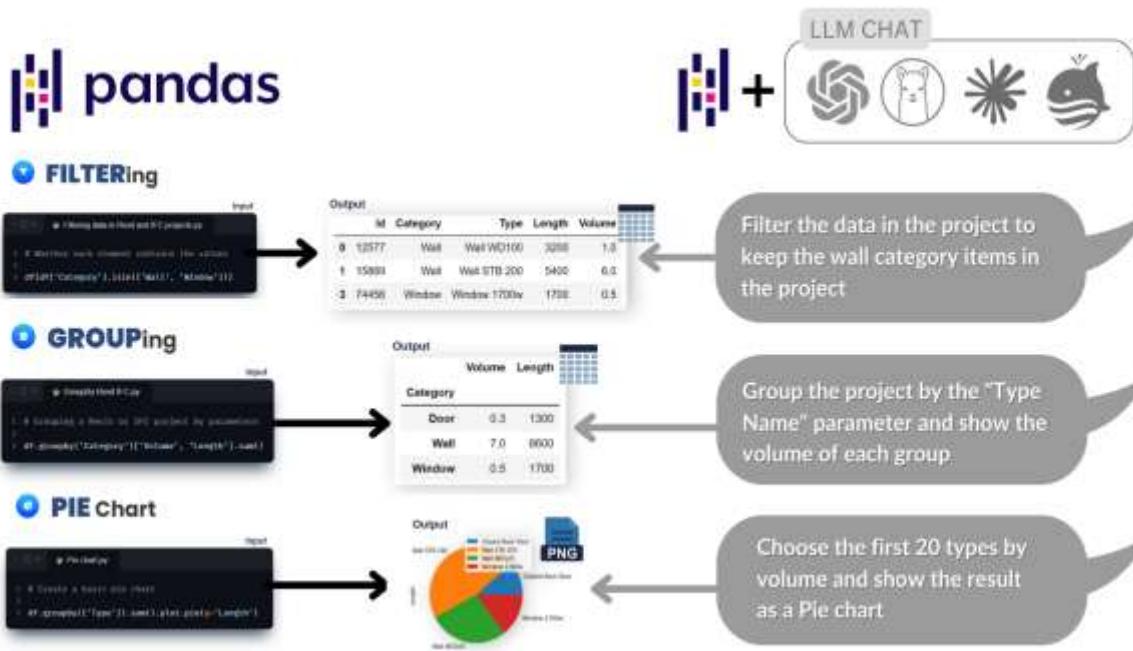
Atšķirībā no CSV un XLSX, Pandas DataFrame nodrošina lielāku elastību un veiktspēju darbā ar datiem: tas var apstrādāt lielu informācijas apjomu operatīvajā atmiņā, atbalsta paplašinātus datu tipus (tostarp datumi, logiskās vērtības un laika rindas) un nodrošina plašas datu filtrēšanas, apkopošanas, apvienošanas un vizualizācijas iespējas. Lai gan CSV neuzkrāj informāciju par datu

tipiem un struktūru, bet XLSX bieži ir pārslogots ar formatēšanu un tam ir zema mērogojamība, DataFrame joprojām ir optimāla izvēle ātrai analīzei, procesu automatizācijai un integrācijai ar mākslīgā intelekta -modeļiem ( 3.4-7. attēls). Turpmākajās nodalās katrs no šiem datu aspektiem tiks aplūkots detalizēti, arī grāmatas 8. daļā tiks detalizēti apskatīti līdzīgi formāti, piemēram, Parquet, Apache Orc, JSON, Feather, HDF5 un datu noliktavas ( 8.1-2. attēls).

		XLSX	CSV	Pandas DataFrame
	Storage	Tabular	Tabular	Tabular
	Usage	Office tasks, data presentation	Simple data exchange	Data analysis, manipulation
	Compression	Built-in	None	None (in-memory)
	Performance	Low	Medium	High (memory dependent)
	Complexity	High (formatting, styles)	Low	Low
	Data Type Support	Limited	Very limited	Extended
	Scalability	Low	Low	Medium (memory limited)

Attēls 3.4-7 DataFrame ir optimāla izvēle datu apstrādei ar augstu veikspēju un uzlabotu datu tipu atbalstu.

Pandas bibliotēka un DataFrame formāts ir kļuvuši par de facto standartu Python datu analīzē, pateicoties to elastībai, jaudai un lietošanas vienkāršībai. Tie ir ideāli piemēroti gan vienkāršu pārskatu veidošanai, gan sarežģītu analītisko konveijeru veidošanai, īpaši kopā ar LLM modeļiem.



Attēls 3.4-8 LLM vienkāršo mijiedarbību ar Pandas: koda vietā pietiek ar teksta vaicājumu.

Mūsdienās Pandas aktīvi izmanto LLM balstītās tērzēšanas istabās, piemēram, ChatGPT, LLaMa, DeepSeek, QWEN un citās. Daudzos gadījumos, kad modelis saņem pieprasījumu, kas saistīts ar tabulas apstrādi, datu validāciju vai analīzi, tas ġenerē kodu, izmantojot tieši Pandas bibliotēku. Tas padara DataFrame par dabisku "valodu" datu attēlošanai dialogos ar mākslīgo intelektuālo intelektu (3.4-8. attēls).

Modernās datu tehnoloģijas, piemēram, Pandas, atvieglo datu analīzi, automatizāciju un integrāciju biznesa procesos. Tās ātri sniedz rezultātus, samazina speciālistu darba slodzi un nodrošina darbību atkārtojamību.

## Nākamie soļi: ilgtspējīgas datu sistēmas izveide

Šajā daļā mēs aplūkojām galvenos būvniecības nozarē izmantotos datu veidus, uzzinājām par dažādiem datu glabāšanas formātiem un analizējām moderno rīku, tostarp LLM un IDE, lomu informācijas apstrādē. Uzzinājām, ka efektīva datu pārvaldība ir pamats pamatotu lēmumu pieņemšanai un uzņēmējdarbības procesu automatizācijai. Organizācijas, kas spēj strukturēt un organizēt savus datus, iegūst ievērojamas konkurences priekšrocības datu apstrādes un pārveidošanas posmos.

Apkopojoš šo daļu, ir vērts izcelt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās pieejas ikdienas uzdevumos:

- Veiciet savu procesu datu auditu
- Veiciet visu projektos izmantoto datu tipu uzskaiti.

- Nosakiet, kuri datu tipi un modeļi ir vissvarīgākie jūsu uzņēmējdarbības procesiem.
- Identificēt problemātiskās jomas, kurās informācija bieži vien ir nestukturēta, slikti strukturēta vai nepieejama.

**Sāciet veidot datu pārvaldības stratēģiju**

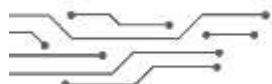
- Izvirzīt politikas jautājumus un standartus, kā rīkoties ar dažādiem datu veidiem.
- Analizējet, kuras no jūsu darba plūsmām var uzlabot, pārvēršot nestukturētus datus strukturētos datos.
- Izveidot datu glabāšanas un piekļuves politiku, kurā nemta vērā drošība un konfidencialitāte.

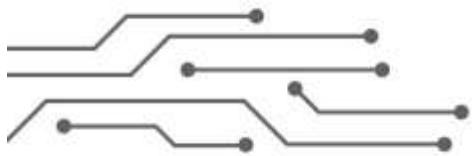
**Instalēt un apgūt pamata rīkus darbam ar datiem**

- Izvēlieties piemērotu IDE, kas atbilst jūsu uzdevumiem (piemēram, instalējet VS Code vai Jupyter Notebook).
- Mēģiniet izveidot vietējo LLM, lai konfidenciāli apstrādātu jūsu personas datus.
- Sāciet eksperimentēt ar Pandas bibliotēku, lai apstrādātu XLSX tabulas datus.
- Aprakstiet LLM tipiskus uzdevumus, ko veicat ar izklājlapi rīkiem vai datubāzēm, un lūdziet LLM automatizēt darbu, izmantojot Pandas.

Šo soļu piemērošana ļaus jums pakāpeniski pārveidot savu pieeju darbam ar datiem, pārejot no atšķirīgām, nestukturētām datu kopām uz vienotu ekosistēmu, kurā dati klūst par pieejamu un saprotamu aktīvu. Sāciet ar maziem darbiem - izveidojiet savu pirmo DataFrame programmā Pandas, palaidiet vietējo LLM, automatizējiet savu pirmo rutīnas uzdevumu, izmantojot Python (piemēram, Excel izklājlapas).

Grāmatas ceturtajā daļā uzmanība tiks pievērsta datu kvalitātei, datu organizēšanai, strukturēšanai un modelēšanai. Mēs pievērsīsimies metodoloģijām, kas pārveido atšķirīgus datu avotus - no PDF un tekstiem līdz attēliem un CAD modeļiem - strukturētās datu kopās, kas piemērotas analīzei un automatizācijai. Mēs arī izpētīsim, kā tiek formalizētas datu prasības, kā būvniecības projektos tiek veidoti konceptuālie un loģiskie modeļi un kā šajā procesā var palīdzēt mūsdienu valodas modeļi (LLM).





## IV DAĻA

### DATU KVALITĀTE: ORGANIZĀCIJA, STRUKTURĒŠANA, MODELĒŠANA.

Ceturtajā daļā galvenā uzmanība pievērsta metodoloģijām un tehnoloģijām, kas paredzētas atšķirīgas informācijas pārveidošanai augstas kvalitātes struktūrētu datu kopās. Detalizēti tiek aplūkoti datu prasību veidošanas un dokumentēšanas procesi, kas ir efektīvas informācijas arhitektūras pamatā būvniecības projektos. Tieki parādītas praktiskas metodes strukturētas informācijas iegūšanai no dažādiem avotiem (PDF -dokumenti, attēli, teksta faili, CAD -modeļi) ar īstenošanas piemēriem. Tieki analizēta regulāro izteiksmju (RegEx) un citu rīku izmantošana automātiskai datu validācijai un verifikācijai. Pakāpeniski aprakstīts datu modelēšanas process konceptuālā, logiskā un fiziskā līmenī, nesmot vērā būvniecības nozares specifiku. Tieki demonstrēti konkrēti piemēri valodas modeļu (LLM) izmantošanai, lai automatizētu informācijas strukturēšanas un validēšanas procesus. Tieki piedāvātas efektīvas pieejas analīzes rezultātu vizualizācijai, palielinot analītiskās informācijas pieejamību visos būvniecības projektu vadības līmeņos.

## NODĀLA 4.1.

### DATU KONVERTĒŠANA STRUKTURĒTĀ FORMĀ

Uz datiem balstītas ekonomikas laikmetā dati kļūst par lēmumu pieņemšanas pamatu, nevis šķērsli. Tā vietā, lai pastāvīgi pielāgotu informāciju katrai jaunai sistēmai un tās formātiem, uzņēmumi arvien biežāk censās izveidot vienotu strukturētu datu modeli, kas kalpo kā universāls patiesības avots visiem procesiem. Mūsdienu informācijas sistēmas tiek veidotas nevis ap formātiem un saskarnēm, bet gan ap datu nozīmi - jo struktūra var mainīties, bet informācijas nozīme paliek nemainīga daudz ilgāk.

Galvenais, lai efektīvi strādātu ar datiem, ir nevis tos bezgalīgi konvertēt un pārveidot, bet gan pareizi organizēt jau no paša sākuma: izveidot universālu struktūru, kas spēj nodrošināt pārredzamību, automatizāciju un integrāciju visos projekta dzīves cikla posmos.

Tradicionālā pieeja liek manuāli veikt pielāgojumus, ieviešot katru jaunu platformu: migrēt datus, mainīt atribūtu nosaukumus un pielāgot formātus. Šie soļi neuzlabo datu kvalitāti, bet tikai maskē problēmas, radot apburto bezgalīgu transformāciju loku. Rezultātā uzņēmumi kļūst atkarīgi no konkrētiem programmatūras risinājumiem, un digitālā transformācija palēninās.

Turpmākajās nodaļās aplūkosim, kā pareizi strukturēt datus un kā izveidot universālus modeļus, līdz minimumam samazināt atkarību no platformas un koncentrēties uz to, kas ir vissvarīgākais - datiem kā stratēģisku resursu, ap kuru tiek veidoti ilgtspējīgi procesi.

### **Mācīšanās, kā pārvērst dokumentus, PDF, attēlus un tekstu strukturētos formātos.**

Būvniecības projektos lielākā daļa informācijas ir nestukturētā veidā: tehniskie dokumenti, darba uzdevumi, rasējumi, specifikācijas, grafiki, protokoli. To dažādība - gan formāta, gan satura ziņā - apgrūtina integrāciju un automatizāciju.

Konvertēšanas process strukturētos vai daļēji strukturētos formātos var atšķirties atkarībā no ievades datu veida un vēlamajiem apstrādes rezultātiem.

Datu pārveidošana no nestukturētas uz strukturētu formu ir gan māksla, gan zinātne. Šis process atšķiras atkarībā no ievades datu veida un analīzes mērķa, un bieži vien tas aizņem nozīmīgu datu inženiera (3.2-5. attēls) un analītiķa (3.2-5. attēls) darba daļu, lai iegūtu tīru, sakārtotu datu kopu.



Attēls 4.1-1 Nestukturēta skenēta dokumenta pārvēršana strukturētas tabulas formātā.

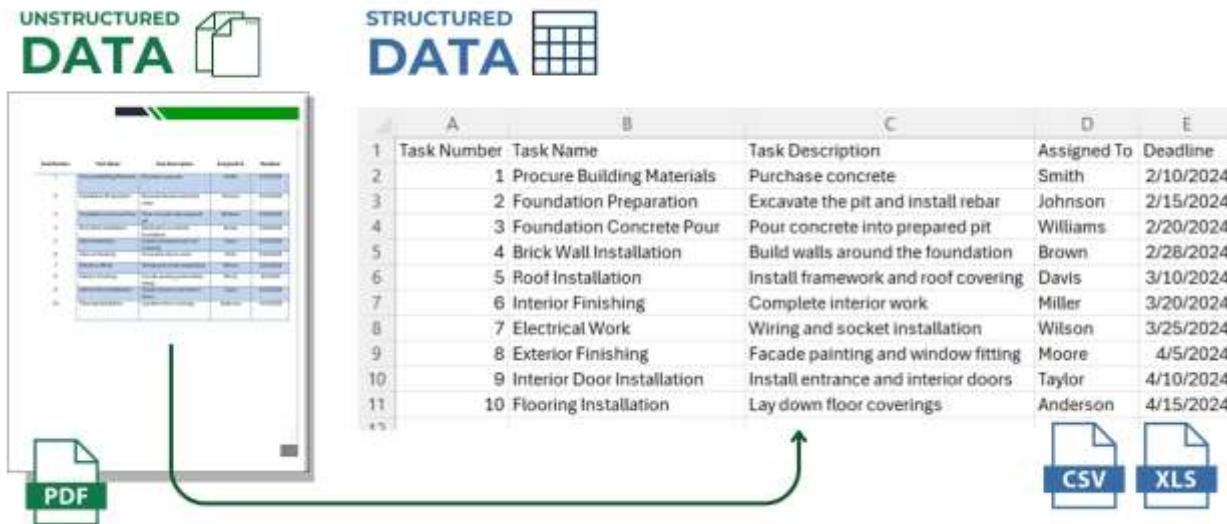
Dokumentu, PDF, attēlu un tekstu pārvēršana strukturētā formātā ( 4.1-1. attēls) ir pakāpenisks process, kas ietver šādas darbības:

- **Izraksts**): Šajā posmā tiek ielādēts avota dokuments vai attēls, kas satur nestukturētus datus. Tas var būt, piemēram, PDF -dokuments, fotogrāfija, rasējums vai shēma.
- **Datu pārveidošana (Transform)**: Pēc tam seko nestukturētu datu pārveidošana strukturētā formātā. Piemēram, tas var ietvert teksta atpazīšanu un interpretēšanu no attēliem, izmantojot optisko rakstzīmju atpazīšanu (OCR) vai citas apstrādes metodes.
- **Datu ielādēšana un saglabāšana (Load)**: pēdējā posmā apstrādātie dati tiek saglabāti dažādos formātos, piemēram, CSV, XLSX, XML, JSON, turpmākajam darbam, kur formāta izvēle ir atkarīga no konkrētām prasībām un vēlmēm.

Šim procesam, kas pazīstams kā ETL (Extract, Transform, Load), ir būtiska nozīme automatizētā datu apstrādē, un tas sīkāk tiks aplūkots nodaļā "ETL un cauruļvads: Extract, Transform, Load". Tālāk aplūkosim piemērus, kā dažādu formātu dokumenti tiek pārveidoti strukturētos datos.

### Piemērs, kā pārvērst PDF -dokumentu tabulā

Viens no biežāk sastopamajiem uzdevumiem būvniecības projektos ir specifikāciju apstrāde PDF formātā. Lai demonstrētu pāreju no nestukturētiem datiem uz strukturētu formātu, aplūkosim praktisku piemēru: tabulas iegūšana no PDF dokumenta un tās konvertēšana CSV vai Excel formātā ( 4.1-2. attēls).



**Attēls. 4.1-2 Atšķirībā no PDF, CSV un XLSX formāti ir plaši izplatīti un viegli integrējami dažādās datu pārvaldības sistēmās.**

LLM valodas modeļi, piemēram, ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN ievērojami vienkāršo datu zinātnieku darbu ar datiem, samazinot nepieciešamību pēc padzīlinātas programmēšanas valodu apguves un ļaujot daudzus uzdevumus risināt ar teksta vaicājumiem.

Tāpēc tā vietā, lai pavadītu laiku, meklējot risinājumus internetā (parasti StackOverflow tīmekļa vietnē vai tematiskajos forumos un tērzētavās) vai sazinoties ar datu apstrādes speciālistiem, mēs varam izmantot mūsdienu tiešsaistes vai vietējo LLM iespējas. Pietiek uzdot vaicājumu, un modelis sniegs gatavu kodu PDF -dokumenta pārvēršanai tabulas formātā.

- 💡 Nosūtiet šādu teksta pieprasījumu jebkuram LLM modelim (CHATGRT, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai citam):

Lūdzu, uzrakstiet kodu, lai iegūtu tekstu no PDF -faila, kurā ir tabula. Šim kodam kā argumentu jāpieņem faila ceļš un jāizraksta izvilkta tabula kā DataFrame ↴

- LLM -modeļa atbilde vairumā gadījumu tiks sniepta kā kods Python valodā , jo šī valoda tiek plaši izmantota datu apstrādei, automatizācijai un darbam ar dažādiem failu formātiem:

```

1 import fitz # PyMuPDF
2 import pandas as pd
3 import re
4
5 # Define the path to the PDF file
6 pdf_path = r'C\Task Number.pdf'
7
8 # Step 1: Extracting text from the PDF
9 def extract_text(pdf_path):
10     """
11     Extract text from a PDF file.
12     Args: pdf_path (str): Path to the PDF file.
13     Returns: str: All text extracted from the PDF.
14     """
15     document = fitz.open(pdf_path)
16     text = ''
17     for page in document:
18         text += page.get_text()
19     return text
20
21 extracted_text = extract_text(pdf_path)

```

Run in IDE

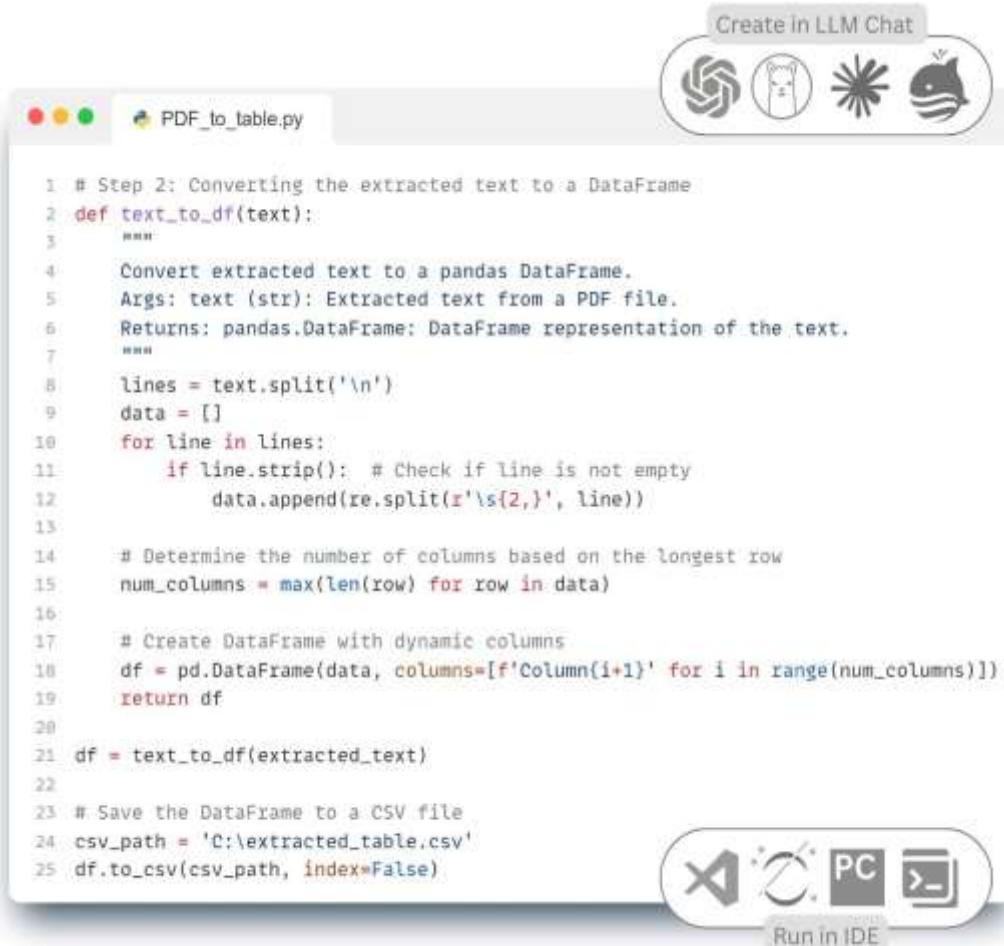
**Attēls 4.1-3 LLM atbilde Python koda veidā un tās bibliotēkas un paketes (Pandas, Fitz) izraksta tekstu no PDF -faila.**

Šo kodu ( 4.1-3. attēls) var palaist bezsaistē vienā no populārākajām IDE, ko minējām iepriekš: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populāros tiešsaistes rīkus: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

- Solī "Konvertēt" mēs izmantojam populāro Pandas bibliotēku (par kuru sīkāk runājām nodalā "Python Pandas: neaizstājams rīks darbam ar datiem"), lai nolasītu iegūto tekstu DataFrame un saglabātu DataFrame CSV tabulas failā vai XLXS:

Man ir nepieciešams kods, kas konvertēs iegūto tabulu no PDF -faila uz DataFrame. Pievienojiet arī kodu, lai saglabātu DataFrame failu CSV failā. ↵

2 LLM atbilde :



```

1 # Step 2: Converting the extracted text to a DataFrame
2 def text_to_df(text):
3     """
4         Convert extracted text to a pandas DataFrame.
5         Args: text (str): Extracted text from a PDF file.
6         Returns: pandas.DataFrame: DataFrame representation of the text.
7     """
8     lines = text.split('\n')
9     data = []
10    for line in lines:
11        if line.strip(): # Check if line is not empty
12            data.append(re.split(r'\s{2,}', line))
13
14    # Determine the number of columns based on the longest row
15    num_columns = max(len(row) for row in data)
16
17    # Create DataFrame with dynamic columns
18    df = pd.DataFrame(data, columns=[f'Column{i+1}' for i in range(num_columns)])
19    return df
20
21 df = text_to_df(extracted_text)
22
23 # Save the DataFrame to a CSV file
24 csv_path = 'C:\extracted_table.csv'
25 df.to_csv(csv_path, index=False)

```

Run in IDE

Attēls 4.1-4 Izvadītās tabulas konvertēšana no PDF uz DataFrame un tabulas saglabāšana CSV failā.

Ja, izpildot kodu, rodas kļūda ( 4.1-3. attēls, 4.1-4. attēls), piemēram, trūkstošu bibliotēku vai nepareiza faila ceļa dēļ, kļūdas tekstu var vienkārši kopēt kopā ar pirmkodu un atkārtoti iesniegt LLM -modelī. Modelis analizēs kļūdas ziņojumu, paskaidros, kāda ir problēma, un ieteiks labojumus vai papildu pasākumus.

Tādējādi mijiedarbība ar AI LLM kļūst par pilnīgu ciklu: p→→pārbaude→ atgriezeniskā saite→ korekcija - bez nepieciešamības pēc padziļinātām tehniskām zināšanām.

Izmantojot vienkārša teksta vaicājumu LLM tērzēšanā un duci rindu Python, ko varam palaist lokālī jebkurā IDE, mēs konvertējām PDF dokumentu tabulārajā CSV formātā, kas atšķirībā no PDF dokumenta ir viegli mašīnlasāms un ātri integrējams jebkurā datu pārvaldības sistēmā.

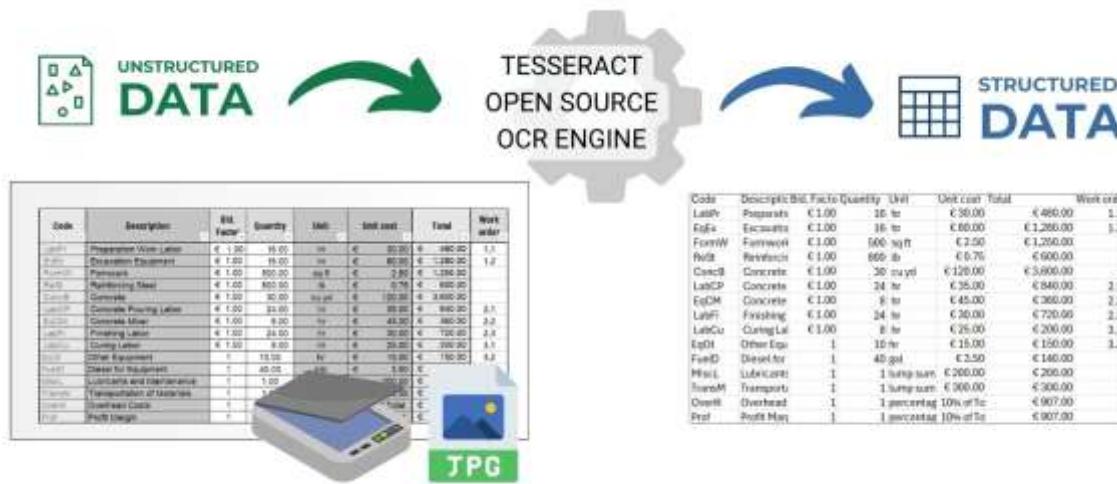
Šo kodu ( 4.1-3. attēls, 4.1-4. attēls), nokopējot to no jebkuras LLM tērzēšanas telpas, mēs varam izmantot desmitiem vai tūkstošiem jaunu PDF dokumentu serverī, tādējādi automatizējot nestrukturētu dokumentu plūsmas pārveidošanas procesu strukturētas CSV tabulas formātā.

Taču PDF dokumentos ne vienmēr ir teksts, visbiežāk tie ir skenēti dokumenti, kas jāapstrādā kā attēli. Lai gan attēli pēc būtības ir nestrukturēti, atpazīšanas bibliotēku izstrāde un izmantošana ļauj iegūt, apstrādāt un analizēt to saturu, ļaujot mums pilnībā izmantot šos datus uzņēmējdarbības procesos.

## Konvertēt JPEG, PNG attēlu uz strukturētu formu

Attēli ir viens no visizplatītākajiem nestrukturētu datu veidiem. Būvniecībā un daudzās citās nozarēs milzīgs informācijas apjoms tiek glabāts skenētu dokumentu, shēmu, fotogrāfiju un rasējumu veidā. Šādi dati satur vērtīgu informāciju, bet tos nevar tieši apstrādāt, piemēram, kā Excel izklājlapu vai datubāzi. Attēli satur daudz sarežģītas informācijas, jo to saturs, krāsas, tekstūras ir daudzveidīgas, un, lai iegūtu noderīgu informāciju, ir nepieciešama īpaša apstrāde.

Grūtības, izmantojot attēlus kā datu avotu, ir struktūras trūkums. Attēli nenodod nozīmi tiešā, viegli izmērāma veidā, ko dators var uzreiz saprast vai apstrādāt, kā to dara Excel izklājlapa vai datubāzes tabula. Lai nestrukturētu attēlu datus pārvērstu strukturētā formā, jāizmanto īpašas bibliotēkas, kas spēj interpretēt tajos ietverto vizuālo informāciju ( 4.1-5. attēls).



Attēls 4.1-5 Skenētu dokumentu un attēlu konvertēšana strukturētos formātos ir iespējama, izmantojot īpašus OCR rīkus.

OCR (optiskā rakstzīmju atpazīšana) tehnoloģija tiek izmantota, lai no attēliem iegūtu tekstu. Tā ļauj atpazīt burtus un ciparus dokumentu skenējumos, fotoattēlos un PDF failos, pārvēršot tos rediģējamā un mašīnlasāmā tekstā. OCR tehnoloģijas jau sen tiek izmantotas dokumentu automatizācijā, un

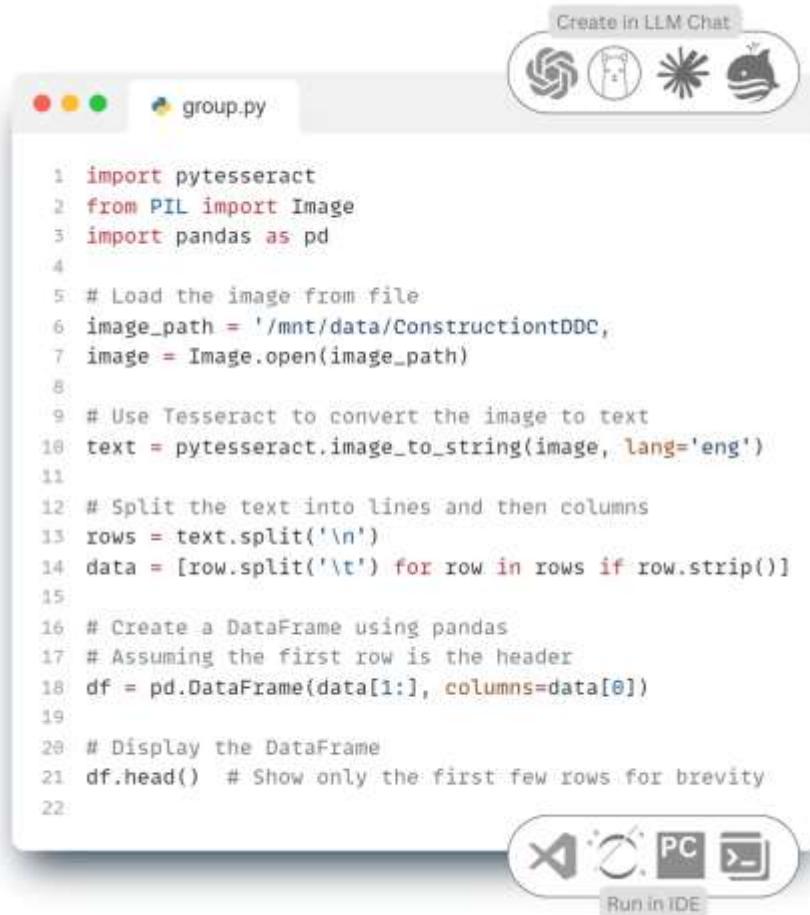
mūsdienās tās ir viegli integrējamas jebkuros biznesa procesos un Python -aplikācijās. Viens no populārākajiem OCR rīkiem ir Tesseract, atvērtā koda, ko sākotnēji izstrādāja HP™ un tagad atbalsta Google™. Tas atbalsta vairāk nekā 100 valodu, un tam ir augsta atpazīšanas precizitāte.

Palūgsim LLM tērzēt, lai uzrakstītu parauga kodu, lai strukturētā veidā iegūtu datus no skenētas vainofotografētas tabulas.

- ⌚ Nosūtiet teksta pieprasījumu LLM tērzēšanai (ChatGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkuru citu):

Uzrakstiet kodu, lai konvertētu JPEG attēlu, kas satur tabulu, uz DataFrame tabulu ↴

- ⌚ LLM atbildē vairumā gadījumu tiek ieteikts izmantot Pytesseract bibliotēku, lai atpazītu tekstu attēlos :



```

1 import pytesseract
2 from PIL import Image
3 import pandas as pd
4
5 # Load the image from file
6 image_path = '/mnt/data/ConstructionDDC'
7 image = Image.open(image_path)
8
9 # Use Tesseract to convert the image to text
10 text = pytesseract.image_to_string(image, lang='eng')
11
12 # Split the text into lines and then columns
13 rows = text.split('\n')
14 data = [row.split('\t') for row in rows if row.strip()]
15
16 # Create a DataFrame using pandas
17 # Assuming the first row is the header
18 df = pd.DataFrame(data[1:], columns=data[0])
19
20 # Display the DataFrame
21 df.head() # Show only the first few rows for brevity
22

```

The screenshot shows a code editor window with a Python script named 'group.py'. The script uses the Pytesseract library to extract text from a JPEG image and then processes it into a pandas DataFrame. The code editor interface includes tabs for 'Create in LLM Chat' and 'Run in IDE'.

Attēls 4.1-6 No attēlu vai fotoattēlu tabulas iegūtā teksta konvertēšana strukturētā tabulārajā attēlojumā.

Šajā piemērā - kods (4.1.1-6. attēls), kas atvasināts no LLM, izmanto pytesseract bibliotēku (Tesseract

for Python), lai pārvērstu attēlu tekstā, izmantojot OCR (optisko rakstzīmju atpazīšanu), un Pandas bibliotēku, lai pārvērstu šo tekstu strukturētā formā, t. i., DataFrame.

Konvertēšanas process parasti ietver pirmapstrādi, lai uzlabotu attēla kvalitāti, pēc tam tiek izmantoti dažādi algoritmi modeļu noteikšanai, pazīmju iegūšanai vai objektu atpazīšanai. Rezultātā nestrukturēta vizuālā informācija tiek pārvērsta strukturētos datos.

Lai gan PDF un attēli ir galvenie nestrukturētas informācijas avoti, patiesais čempions apjoma ziņā ir teksts, kas tiek ģenerēts e-pasta ziņojumos, tērzētavās, sanāksmēs, ziņojumapmaiņas ziņojumos. Šo datu ir ne tikai daudz - tie ir izkliedēti, neformalizēti un ārkārtīgi slikti strukturēti.

## Teksta datu pārveidošana strukturētā formā

Papildus PDF dokumentiem ar tabulām (4.1-2. attēls) un skenētām tabulu veidlapu versijām (4.1-5. attēls) ievērojama daļa informācijas projekta dokumentācijā tiek sniepta teksta veidā. Tā var būt gan sakarīgi teikumi teksta dokumentos, gan fragmentāri ieraksti, kas izkaisīti pa rasējumiem un shēmām. Mūsdienu datu apstrādes apstākļos viens no biežāk sastopamajiem uzdevumiem ir pārveidot šādu tekstu strukturētā formātā, kas piemērots analīzei, vizualizācijai un lēmumu pieņemšanai.

Šī procesa centrā ir taksonomija - klasifikācijas sistēma, kas ļauj informāciju sakārtot kategorijās un apakškategorijās, pamatojoties uz kopīgām iezīmēm.

**Taksonomija** ir hierarhiska klasifikācijas struktūra, ko izmanto objektu grupēšanai un organizēšanai. Teksta apstrādes kontekstā tā kalpo par pamatu sistemātiskai objektu iedalīšanai semantiskās kategorijās, tādējādi vienkāršojot analīzi un uzlabojot datu apstrādes kvalitāti.

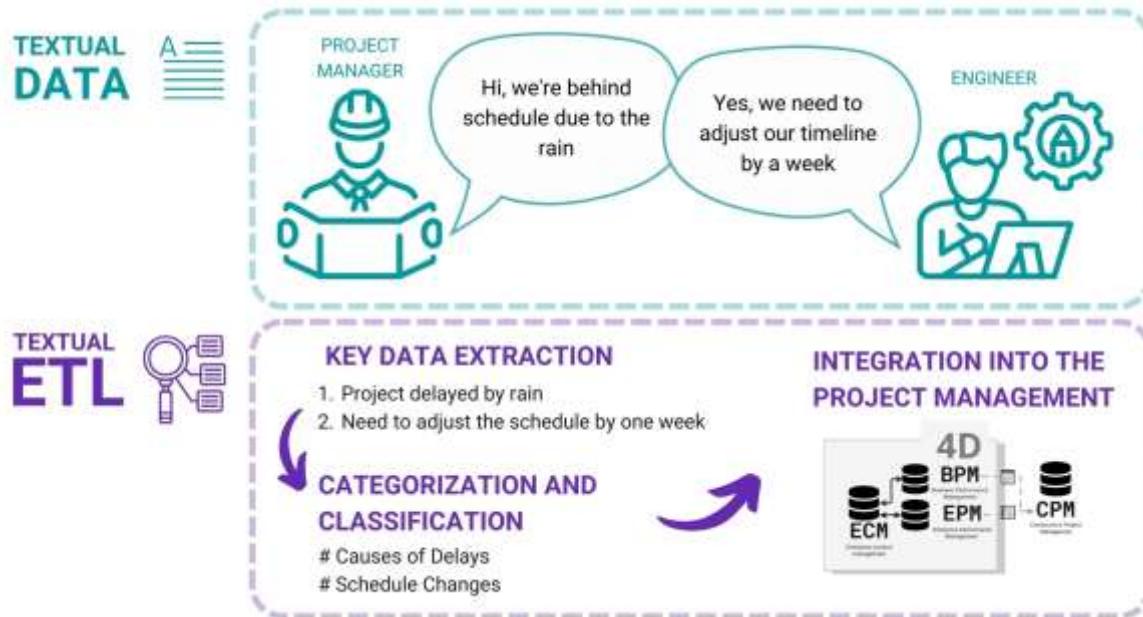
Taksonomijas izveide tiek papildināta ar būtību ieguves, kategorizēšanas un kontekstualizācijas posmiem. Lai modelētu informācijas ieguves procesu no teksta datiem, mums ir jāveic šādi soli, kas ir līdzīgi tiem, kurus jau esam piemērojuši datu strukturēšanai no PDF dokumentiem:

- Izraksts): jums jāanalizē teksta dati, lai iegūtu informāciju par kavējumiem un izmaiņām projekta grafikā.
- **Kategorizēšana un klasificēšana** (Transofrm): kategorizē saņemto informāciju, piemēram, kavējumu un grafika izmaiņu iemeslus.
- **Integrācija** (Load): beigās mēs sagatavojam strukturētus datus integrācijai ārējās datu pārvaldības sistēmās.

Apskatiet situāciju: mums ir dialogs starp projekta vadītāju un inženieri, kurā tiek apspriestas problēmas ar grafiku. Mūsu mērķis ir izcelt galvenos elementus (kavēšanās iemeslus, grafika korekcijas) un strukturēti tos izklāstīt (4.1-7. attēls).

Veiksim ieguvi, pamatojoties uz paredzamajiem atslēgvārdiem, izveidosim DataFrame, lai simulētu datu ieguvi, un pēc pārveidošanas izveidosim jaunu DataFrame tabulu, kurā būs datuma, notikuma

(piemēram, kavēšanās iemesls) un darbības (piemēram, grafika izmaiņas) kolonnas.



Attēls 4.1-7 Izcelta galvenā informācija no teksta par nepieciešamību pielāgot termiņus un integrēt izmaiņas projekta vadības sistēmā.

Šeit ir kods, lai atrisinātu problēmu, izmantojot teksta vaicājumu vienā no valodas modeļiem, kā iepriekšējos piemēros.

❷ Nosūtiet teksta pieprasījumu uz jebkuru LLM tērzēšanas istabu:

Man ir saruna starp vadītāju: "Labdien, lietus dēļ mēs atpaliekam no grafika" un inženieri: "Jā, mums ir jāpielāgo termiņš par nedēļu". Man ir nepieciešams skripts, kas analizētu turpmākus līdzīgus teksta dialogus, izvilktu no tiem kavējumu iemeslus un nepieciešamās termiņu korekcijas un pēc tam no šiem datiem ģenerētu DataFrame. DataFrame pēc tam ir jāieraksta CSV failā. ↴

- LLM atbilde parasti ietver Python kodu, izmantojot regulārās izteiksmes (re - Regex) un Pandas (pd) bibliotēku :

```

import pandas as pd
import re

# Sample text data from a discussion between a manager and an engineer.
text_data = [
    "Hi, we're behind schedule due to the rain",
    "Yes, we need to adjust our timeline by a week".
]

# Extract key information elements
def extract_key_elements(text):
    delay_pattern = r"(behind schedule|delayed) due to the (\w+)"
    time_adjust_pattern = r"adjust our timeline by a (\w+)"

    delay_search = re.search(delay_pattern, text)
    time_adjust_search = re.search(time_adjust_pattern, text)

    delay_reason = delay_search.group(2) if delay_search else None
    time_adjustment = time_adjust_search.group(1)
        if time_adjust_search else None

    return delay_reason, time_adjustment

# Applying the function to our text data
extracted_data = [extract_key_elements(text) for text in text_data]

# Convert the data into a DataFrame
df_extracted = pd.DataFrame(extracted_data,
columns=['Cause of Delay', 'Time Adjustment'])

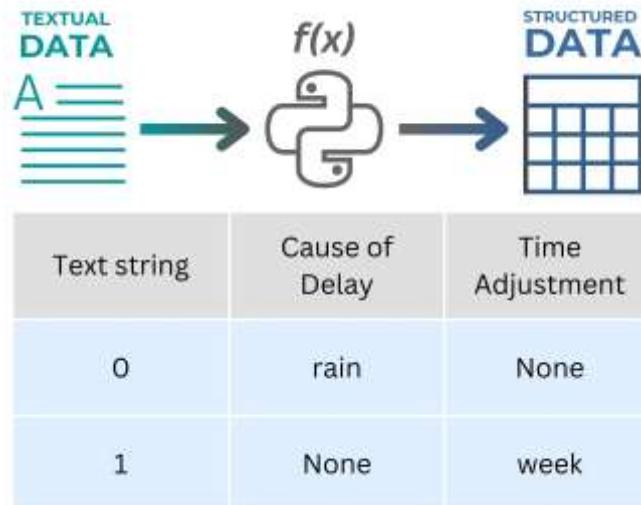
```

Attēls 4.1-8. Galvenās informācijas izcelšana tekstā par nepieciešamību pielāgot termiņus tabulā.

Šajā piemērā ( 4.1-7. attēls) teksta dati, kas satur saraksti starp projekta vadītāju un inženieri, tiek analizēti, lai identificētu un iegūtu specifisku informāciju, kas var ietekmēt turpmāko projektu vadību ar līdzīgiem dialogiem. Izmantojot regulārās izteiksmes (vairāk par regulārajām izteiksmēm mēs aplūkosim nodalā "Strukturētās prasības un RegEx regulārās izteiksmēs "), izmantojot modeļus, tiek identificēti projekta kavējumu cēloņi un nepieciešamās korekcijas laika grafikā. Šajā piemērā uzrakstītā funkcija no virknēm, pamatojoties uz šabloniem, izraksta vai nu kavēšanās cēloni, vai laika korekciju: izvēloties vārdu pēc "jo of" kā kavēšanās cēloni vai vārdu pēc "līdz" kā laika korekciju.

Ja rindā ir minēta kavēšanās laikapstākļu dēļ, kā iemesls tiek norādīts "lietus"; ja rindā ir minēta grafika korekcija konkrētam periodam, šis periods tiek izdalīts kā laika korekcija ( 4.1-9. attēls). Ja rindā nav

neviena no šiem vārdiem, attiecīgā atribūta slejas vērtība ir "Nav".



**Attēlā 4.1-9. attēls Kopsavilkuma tabulā, kas iegūta kā DataFrame, pēc koda izpildes ir informācija par kavējumiem un nepieciešamajām laika korekcijām.**

Nosacījumu strukturēšana un parametrizēšana no teksta (dialoga, vēstules, dokumenta) ļauj ātri novērst kavēšanos būvniecībā: piemēram, strādnieku trūkums var ietekmēt darbu tempu sliktos laika apstākļos, tāpēc uzņēmumi, zinot kavēšanās parametrus no dialogiem (4.1-9. attēls) starp meistarū būvlaukumā un projekta vadītāju - iepriekš var stiprināt brigādi nelabvēlīgas prognozes gadījumā.

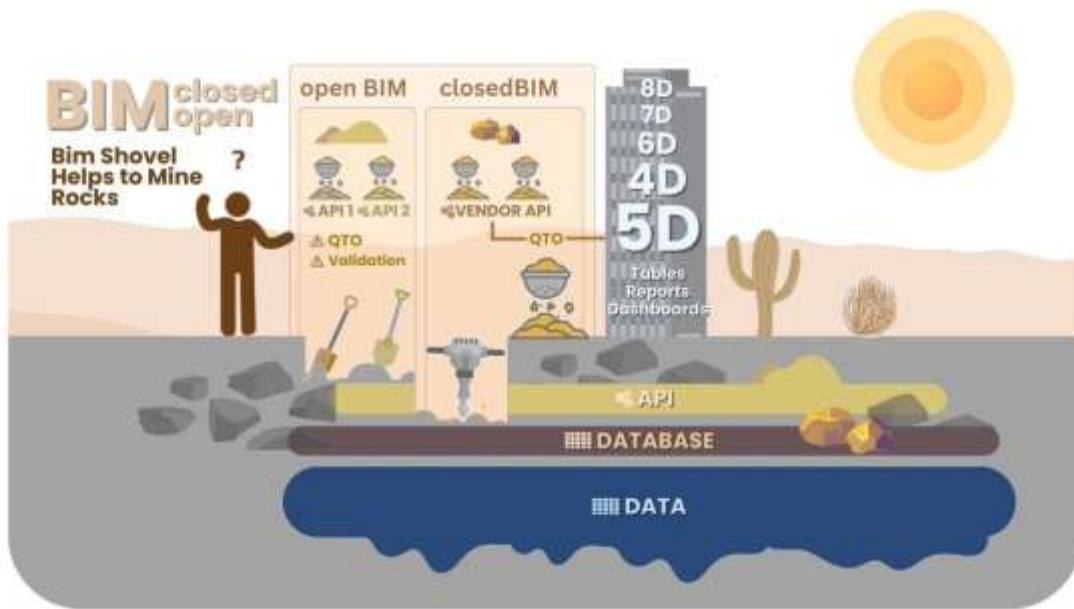
Dokumentu un attēlu pārveidošanu strukturētā formātā var veikt, izmantojot saīdzinoši vienkāršus, atvērtus un bezmaksas uz kategorizāciju balstītus rīkus.

Arī elementu kategorizēšana ir būtiska darba ar projekta datiem daļa, jo īpaši saistībā ar CAD programmatūras izmantošanu (BIM).

## CAD datu (BIM) konvertēšana strukturētā formā

CAD datu strukturēšana un kategorizēšana (BIM) ir sarežģītāks uzdevums, jo CAD (BIM) datubāzēs glabātie dati gandrīz vienmēr ir slēgtos vai sarežģītos parametriskos formātos, bieži vien vienlaikus apvienojot ģeometrisko datu elementus (daļēji strukturēti) un metainformācijas elementus (daļēji strukturēti vai strukturēti dati).

Natīvie datu formāti CAD (BIM) sistēmās parasti ir aizsargāti un nav pieejami tiešai lietošanai, ja vien nav specializētas programmatūras vai API - paša izstrādātāja saskarnes (4.1-10. att.). Šāda datu izolācija veido slēgtus glabāšanas silosus, kas ierobežo brīvu informācijas apmaiņu un kavē visaptverošu digitālo procesu izveidi uzņēmumā.



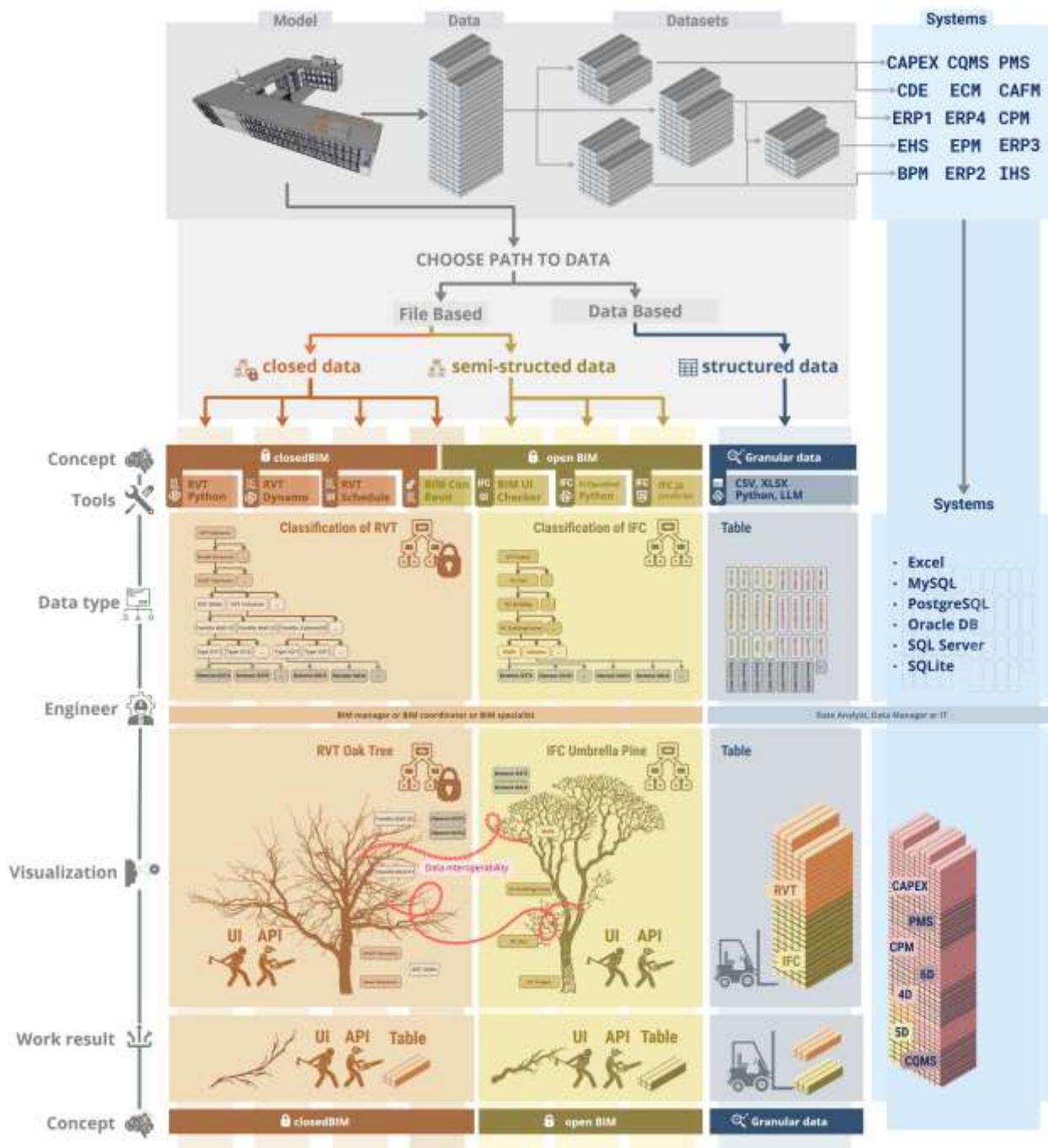
Attēls 4.1-10 CAD speciālisti (BIM) var piekļūt vietējiem datiem, izmantojot API savienojumus vai piegādātāju rīkus.

Speciālos CAD (BIM) formātos informācija par projekta elementu īpašībām un atribūtiem tiek apkopota hierarhiskā klasifikācijas sistēmā, kur vienības ar atbilstošām īpašībām, līdzīgi augļu koka augļiem, atrodas jaunākajos datu klasifikācijas zaru mezglos (4.1-11. attēls).

Datu ieguve no šādām hierarhijām ir iespējama divējādi: vai nu manuāli, noklikšķinot uz katra mezglā, it kā apstrādātu koku, ar cirvi nogriežot atlasītos kategoriju un tipu zarus. Alternatīvi, lietojumprogrammu saskarņu (API) izmantošana nodrošina efektīvāku, automatizētu pieeju datu iegūšanai un grupēšanai, galu galā pārveidojot tos strukturētā tabulā izmantošanai citās sistēmās.

Strukturētu datu tabulu iegūšanai no CAD (BIM) projektiem var izmantot dažādus rīkus, piemēram, Dynamo, pyRvt, Pandamo (Pandas + Dynamo), ACC vai atvērtā koda risinājumus, piemēram, IfcOpSh vai IFCjs IFC formātam.

Mūsdienīgi datu eksportēšanas un konvertēšanas rīki ļauj vienkāršot datu apstrādi un sagatavošanu, sadalot CAD modeļu saturu divās galvenajās sastāvdaļās: ģeometrijas informācija un atribūtu dati (4.1-13. attēls) - metainformācija, kas apraksta konstrukcijas elementu īpašības (3.1-16. attēls). Šie divi datu slāni paliek saistīti ar unikāliem identifikatoriem, pateicoties kuriem katru elementu ar ģeometrijas aprakstu (izmantojot parametrus vai daudzstūrus) ir iespējams precīzi sasaistīt ar tā atribūtiem: nosaukumu, materiālu, pabeigšanas stadiju, izmaksām utt. Šāda pieeja nodrošina modeļa integritāti un ļauj elastīgi izmantot datus gan vizualizācijai (ģeometriskā modeļa dati), gan analītiskiem vai pārvaldības uzdevumiem (strukturēti vai brīvi strukturēti), strādājot ar abiem datu veidiem atsevišķi vai paralēli.



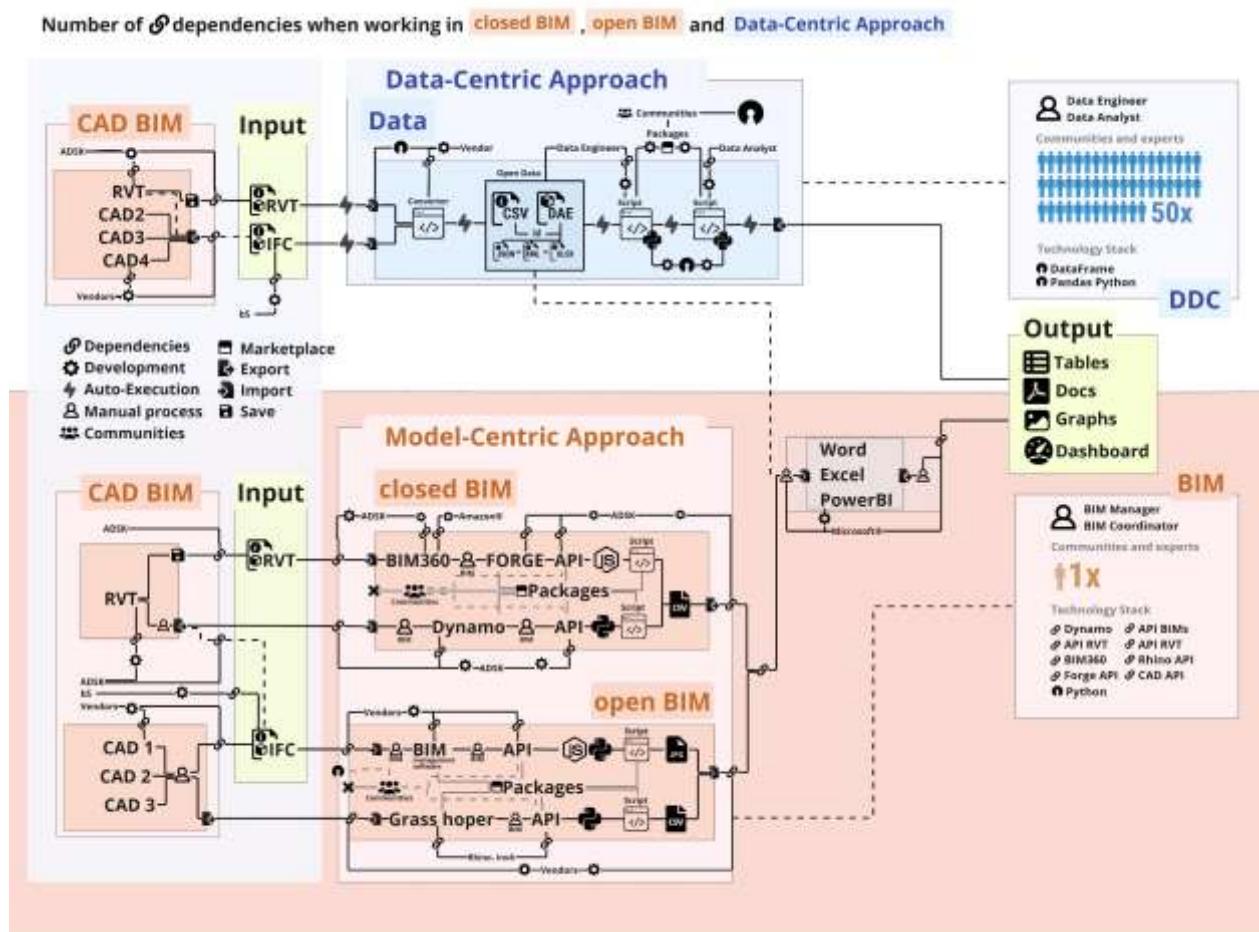
Attēls 4.1-11 CAD datubāzu informācijas skats (BIM) lietotājam tiek attēlots klasifikācijas koku veidā.

Attīstoties reversās inženierijas tehnoloģijām un parādoties SDK (Software Development Kit) CAD datu konvertēšanai, datu pieejamība un konvertēšana no slēgtiem CAD programmatūras formātiem (BIM) ir kļuvusi daudz vienkāršāka. Tagad ir iespējams legāli un droši konvertēt datus no slēgtiem formātiem universālos formātos, kas piemēroti analīzei un izmantošanai citās sistēmās. Pirmo reversās inženierijas rīku ("Open DWG") vēsture un cīņa par dominanci pār CAD ražotāju formātiem tika aplūkota

nodaļā "Strukturētie dati: digitalās transformācijas pamats".

Reversās inženierijas rīki jauj likumīgi iegūt datus no slēgtiem patentētiem formātiem, sadalot informāciju no jaukta CAD (BIM) formāta lietotājam vajadzīgajos datu veidos un formātos, tādējādi atvieglojot tās apstrādi un analīzi.

Izmantojot reverso inženieriju un tiešu piekļuvi informācijai no CAD datubāzēmpadara informāciju pieejamu, jaujot izmantot atvērtus datus un atvērtus rīkus, kā arī datu analīzi, izmantojot standarta rīkus, veidojot pārskatus, vizualizācijas un integrāciju ar citām digitālajām sistēmām ( 4.1-12. att.).



Attēls 4.1-12 Tiešā piekļuve CAD datiem samazina atkarību no programmatūras platformām un veicina uz datiem orientētu pieeju.

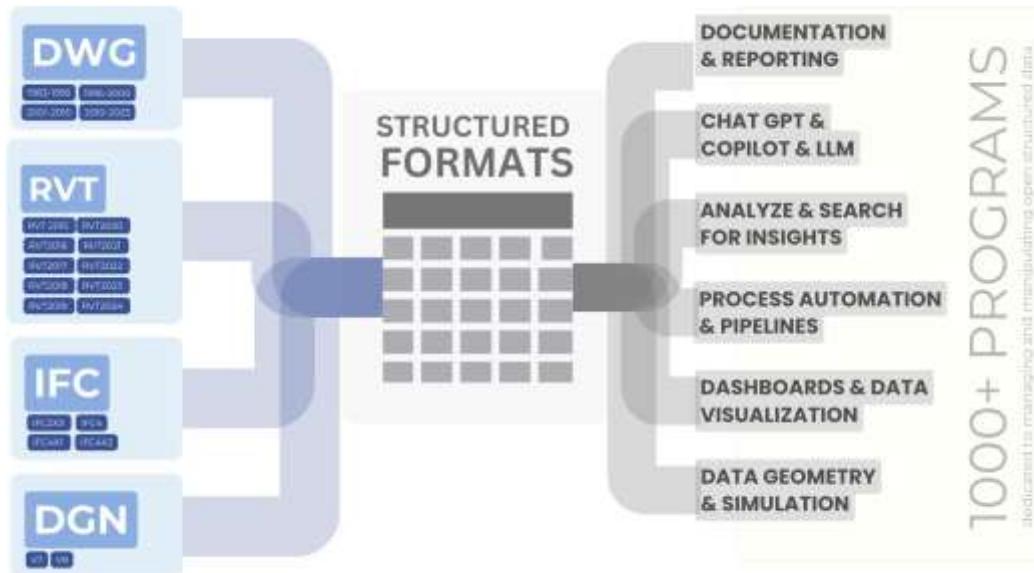
Kopš 1996. gada DWG formātā, kopš 2008. gada DGN formātā un kopš 2018. gada RVT formātā sākotnēji slēgtos CAD datu formātus ir iespējams ērti un efektīvi konvertēt jebkurā citā formātā, tostarp strukturētā formātā, izmantojot reversās inženierijas rīkus ( 4.1-13. attēls). Mūsdienās gandrīz visi lielākie CAD (BIM) un lielie inženiertehniskie uzņēmumi pasaulē izmanto SDK - reversās inženierijas rīkus, lai iegūtu datus no slēgtajiem CAD (BIM) piegādātāju formātiem [92].



**Attēls 4.1-13 Reversās inženierijas rīku izmantošana ļauj CAD (BIM) programmu datubāzes pārvērst jebkurā ērtā datu modelī.**

Datu konvertēšana no slēgtiem, patentētiem formātiem uz atvērtiem formātiem un jauktu CAD (BIM) formātu sadalīšana ģeometriskos un metainformācijas atribūtu datos vienkāršo darbu ar tiem, padarot tos pieejamus analīzei, manipulācijām un integrācijai ar citām sistēmām ( 4.1-14. att.).

Mūsdienās, strādājot ar CAD datiem (BIM), mēs esam sasnieguši punktu, kad mums nav jāprasa atļauja no CAD (BIM) piegādātājiem, lai piekļūtu informācijai no CAD formātiem.



**Attēls 4.1-14 Mūsdienu SDK rīki ļauj legāli konvertēt datus no patentētiem CAD datubāzes formātiem (BIM).**

Pašreizējās CAD projektēšanas datu apstrādes tendences turpina veidot galvenie tirgus dalībnieki - CAD - piegādātāji, kas cenšas nostiprināt savas pozīcijas datu pasaulē un rada jaunus formātus un koncepcijas.

## CAD risinājumu piedāvātāji virzās uz strukturētu datu apstrādi

Sākot ar 2024. gadu, projektēšanas un būvniecības nozarē notiks ievērojamas tehnoloģiskās pārmaiņas datu izmantošanā un apstrādē. CAD -sistēmu ražotāji pievēršas nevis brīvai piekļuvei projektēšanas datiem, bet gan nākamo jauno koncepciju veicināšanai. Tādas pieejas kā BIM (izveidots 2002. gadā) un atvērtais BIM (izveidots 2012. gadā) pakāpeniski piekāpjas mūsdienīgiem tehnoloģiskiem risinājumiem, kurus sāk popularizēt CAD pārdevēji [93]:

- Pāreja uz "granulāru" datu izmantošanu, kas ļauj efektīvi pārvaldīt informāciju un pāriet uz datu analīzi.
- USD formāta izveide un entitāšu-komponentu-sistēmas pieejas (ECS) ieviešana elastīgai datu organizācijai.
- Aktīva mākslīgā intelekta izmantošana datu apstrādē, procesu automatizācijā un datu analīzē.
- Sadarbspējas attīstīšana - uzlabota mijiedarbība starp dažādām programmām, sistēmām un datubāzēm.

Katrs no šiem aspektiem tiks sīkāk aplūkots grāmatas "CAD un BIM: mārketing, realitāte un projektēšanas datu nākotne būvniecībā" sestajā daļā. Šajā nodalā mēs tikai īsi ieskicēsim vispārējo pārmaiņu vektoru: lielākie CAD ražotāji šobrīd cenšas pārskatīt projektēšanas informācijas strukturēšanas veidu. Viena no galvenajām pārmaiņām ir atteikšanās no klasiskā uz failiem balstītā glabāšanas modeļa par labu uz analītiskiem datiem orientētai granulārai datu arhitektūrai, kas nodrošina nepārtrauktu piekļuvi atsevišķām modeļa sastāvdaļām [93].

Notiekosā būtība ir tāda, ka nozare pakāpeniski atsakās no aprūtinošiem, specializētiem un parametriskiem formātiem, kuriem nepieciešami ģeometriskie kodoli, dodot priekšroku universālākiem, mašīnlasāmiem un elastīgākiem risinājumiem.

Viens no šādiem pārmaiņu virzītājspēkiem ir USD (Universal Scene Description) formāts, kas sākotnēji tika izstrādāts datorgrafikas nozarē, bet jau ir atzīts inženiertehniskajās lietojumprogrammās, pateicoties NVIDIA Omniverse (un Isaac Sim) platformai simulačijām un vizualizācijām [93]. Atšķirībā no parametriskā IFC, USD piedāvā vienkāršāku struktūru un ļauj aprakstīt ģeometriju un objektu īpašības JSON formātā (4.1-15. attēls), kas atvieglo informācijas apstrādi un paātrina tās integrāciju digitālajos procesos. Jaunais formāts ļauj uzglabāt ģeometriju (papildus BREP -NURBS - sīkāk grāmatas 6. daļā) MESH daudzstūru formā, bet objektu īpašības - JSON formātā, kas padara to ērtāku automatizētiem procesiem un darbam mākoņu ekosistēmās [94].

Daži CAD un ERP ražotāji jau izmanto līdzīgus formātus (piemēram, NWD, SVF, CP2, CPIXML), taču lielākā daļa no tiem joprojām ir slēgti un nav pieejami ārējai lietošanai, kas ierobežo datu integrācijas un atkārtotas izmantošanas iespējas. Šajā kontekstā USD var spēlēt tādu pašu lomu kā savulaik DXF - atvērta alternatīva tādiem patentētiem formātiem kā DWG.

General Information				Comparison / Notes
Year of format creation	1991	2016		IFC focuses on construction data, USD on 3D graphics
Creator-developer	TU Munich	Pixar		IFC was founded in Germany, USD in America
Prototypes and predecessors	IGES, STEP	PTEX, DAE, GLTF		IFC evolved from IGES/STEP, USD from PTEX/DAE/GLTF
Initiator in Construction	ADSK	ADSK		ADSK initiated the adoption of both formats in construction
Organizer of the Alliance	ADSK	ADSK		ADSK organized both alliances
Name of the Alliance	b5 (IAI)	AOUSD		Different alliances for each format
Year of Alliance Formation	1994	2023		The IFC alliance was formed in 1994, AOUSD for USD in 2023
Promoting in the construction	ADSK and Co	ADSK and Co		ADSK and Co actively promotes both formats in b5 (IAI) since the introduction

Purpose and Usage				Comparison / Notes
Purpose	Semantic description and interoperability	Data simplification, visualization unification		IFC for semantics and exchange; USD for simplification and visualization
Goals and Objectives	Interoperability and semantics	Unification for visualization and data processing		IFC focuses on semantics; USD on visualization
Use in Other Industries	Predominantly in construction	In film, games, VR/AR, and now in construction		USD is versatile and used in various fields
Supported Data Types	Geometry, object attributes, metadata	Geometry, shaders, animation, light, and camera		USD supports a wider range of data types suitable for complex visualizations; IFC focuses on construction-specific data

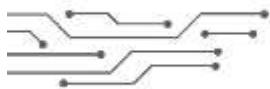
Attēls 4.1-15 USD formāts kā CAD ražotāju mēģinājums apmierināt pieprasījumu pēc savietojamības un projektēšanas datu neatkarības no ģeometriskiem kodoliem.

Lielāko izstrādātāju pāreja uz atvērtiem un vienkāršotiem USD, GLTF, OBJ, XML (slēgtie NWD, CP2, SVF, SVF2, CPIXML) un līdzīgiem formātiem ( 3.1-17. attēls) atspoguļo globālo tendenci un nozares pieprasījumu pēc datu vienkāršošanas un lielākas pieejamības. Turpmākajos gados var sagaidīt pakāpenisku atteikšanos no sarežģītiem parametriskiem standartiem un formātiem, kas atkarīgi no ģeometriskiem kodoliem, par labu vieglākiem un strukturētākiem risinājumiem. Šī pāreja paātrinās būvniecības nozares digitalizāciju, veicinās procesu automatizāciju un vienkāršos datu apmaiņu.

Neraugoties uz CAD -pārdevēju stratēģiskajiem plāniem veicināt jaunus atvērtus formātus, būvniecības nozares profesionāļi var arī pilnībā pieķūt datiem no slēgtām CAD sistēmām, neizmantojot CAD (BIM) rīkus, izmantojot reversās inženierijas rīkus.

Visas šīs tendences neizbēgami novēd pie pārejas no apjomīgiem, monolītiem 3D modeļiem uz universāliem, strukturētiem datiem un citās nozarēs jau sen pārbaudītu formātu izmantošanu. Tiek līdz projekta komandas sāk uztvert CAD modeļus ne tikai kā vizuālus objektus vai datņu kopumu, bet gan kā datu bāzes, kas satur zināšanas un informāciju, pieeja projektēšanai un pārvaldībai krasī mainās.

Kad komandas ir iemācījušās iegūt strukturētus datus no dokumentiem, tekstiemi, rasējumiem un CAD modeļiem un tām ir piekļuve datubāzēm, nākamais galvenais solis ir datu modelēšana un kvalitātes nodrošināšana. Tieši šīs solis lielā mērā nosaka informācijas apstrādes un pārveidošanas ātrumu, kas galu galā tiks izmantots, lai pieņemtu lēmumus konkrētos lietojumu uzdevumos.



## NODAĻA 4.2.

### **KLASIFIKĀCIJA UN INTEGRĀCIJA: VIENOTA BŪVNIECĪBAS DATU VALODA**

#### **Lēmumu pieņemšanas ātrums ir atkarīgs no datu kvalitātes**

Mūsdienu dizaina datu arhitektūrā notiek būtiskas pārmaiņas. Nozarē notiek pāreja no apjomīgiem, izolētiem modeļiem un slēgtiem formātiem uz elastīgākām, mašīnlasāmām struktūrām, kas vērstas uz analīzi, integrāciju un procesu automatizāciju. Tomēr pāreja uz jauniem formātiem vien negarantē efektivitāti - uzmanības centrā neizbēgami ir pašu datu kvalitātei.

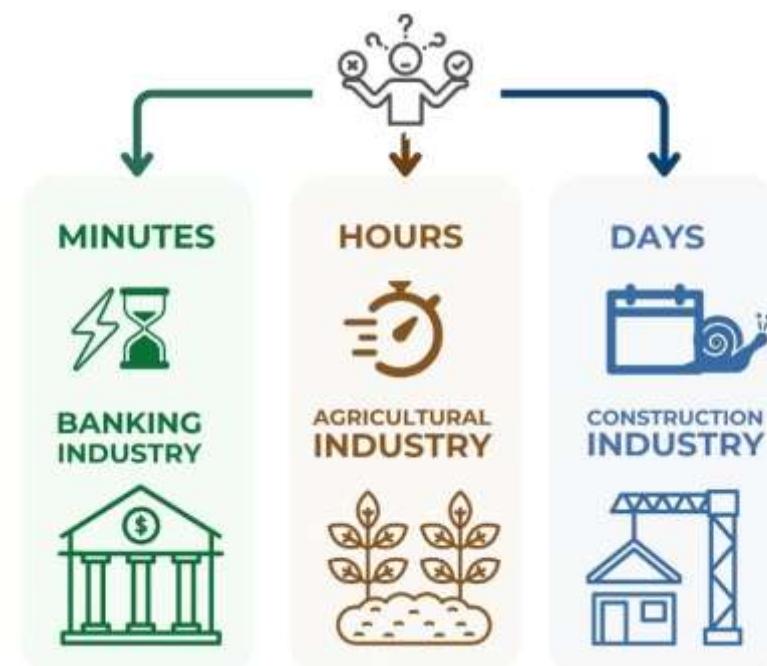
Šīs grāmatas lappusēs mēs daudz runājam par formātiem, sistēmām un procesiem. Taču visi šie centieni ir bezjēdzīgi bez viena svarīga elementa - datiem, kuriem var uzticēties. Datu kvalitāte ir digitalizācijas stūrakmens, un mēs pie tās atgriezīsimies turpmākajās daļās.

Mūsdienu būvniecības uzņēmumi - īpaši lielie - izmanto desmitiem un dažkārt pat tūkstošiem dažādu sistēmu un datubāzu ( 4.2-1. attēls). Šīs sistēmas ne tikai regulāri jāpapildina ar jaunu informāciju, bet arī efektīvi jāsadarbojas savā starpā. Visi jaunie dati, kas rodas ienākošās informācijas apstrādes rezultātā, tiek integrēti šajās vidēs un kalpo konkrētu uzņēmējdarbības uzdevumu risināšanai.

Un, ja agrāk lēmumus par konkrētiem uzņēmējdarbības uzdevumiem pieņēma augstākā līmeņa vadītāji - tā sauktie HiPPO ( 2.1-9. attēls) -, pamatojoties uz pieredzi un intuīciju, tad šodien, strauji pieaugot informācijas apjomam, šāda pieeja klūst pretrunīga. To aizstāj automatizētā analīze, kas strādā ar reāllaika datiem.

"Tradicionāli-manuālās" vadītāju līmeņa uzņēmējdarbības procesu diskusijas pāries uz operatīvo analīzi, kas prasa ātras atbildes uz uzņēmējdarbības pieprasījumiem.

Laiks, kad grāmatveži, meistari un tāmētāji dienām un nedēļām veidoja manuāli ģenerētus pārskatus, kopsavilkuma tabulas un projekta datu vitrīnas, ir pagātne. Mūsdienās lēmumu pieņemšanas ātrums un savlaicīgums klūst par galveno konkurences priekšrocību faktoru.



Attēlā: 4.2-1 Būvniecības nozarē aprēķinu veikšanai un lēmumu pieņemšanai nepieciešamas dienas, atšķirībā no citām nozarēm, kur tas notiek stundās vai minūtēs.

Galvenā atšķirība starp būvniecības nozari un digitāli attīstītākām nozarēm (4.2-1. attēls) zemais datu kvalitātes un standartizācijas līmenis. Novecojušas pieejas informācijas ģenerēšanai, pārraidei un apstrādei palēnina procesus un rada haosu. Vienotu datu kvalitātes standartu trūkums kavē visaptverošas automatizācijas ieviešanu.

Viena no galvenajām problēmām joprojām ir zema ievades datu kvalitāte, kā arī formalizētu procesu trūkums to sagatavošanai un apstiprināšanai. Bez uzticamiem un konsekventiem datiem nav iespējama efektīva sistēmu integrācija. Tas noved pie kavēšanās, klūdām un lielākām izmaksām visos projekta dzīves cikla posmos.

Turpmākajās grāmatas sadaļās detalizēti aplūkosim, kā uzlabot datu kvalitāti, standartizēt procesus un saīsināt ceļu no informācijas līdz kvalitatīviem, apstiprinātiem un konsekventiem datiem.

## Datu standartizācija un integrācija

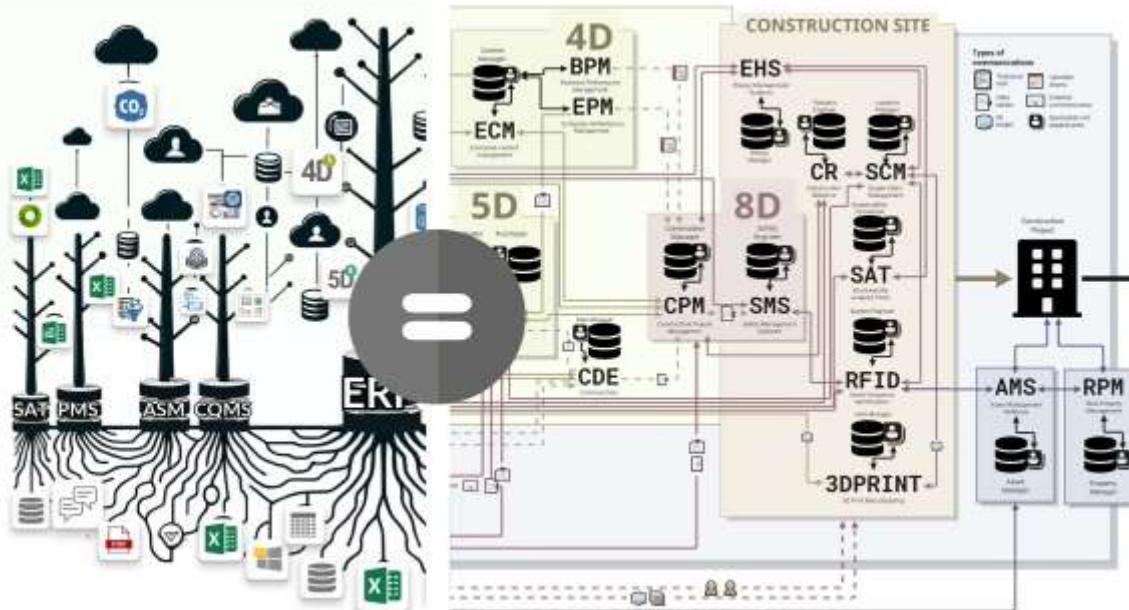
Efektīvai datu pārvaldībai nepieciešama skaidra standartizācijas stratēģija. Tikai ar skaidrām datu struktūras un kvalitātes prasībām var automatizēt datu validāciju, samazināt manuālo darbību skaitu un paātrināt uz informāciju balstītu lēmumu pieņemšanu visos projekta posmos.

Ikdienas praksē būvniecības uzņēmumam katru dienu ir jāapstrādā simtiem failu: e-pasti, PDF dokumenti, CAD dizaina faili, dati no IOT sensoriem, kas jāintegrē uzņēmuma biznesa procesos.

Uzņēmuma datu bāzu un rīku ekosistēmas mežam (4.2-2. attēls) ir jāiemācās iegūt barības vielas no

ienākošajiem daudzformāta datiem, lai sasniegtu uzņēmumam vēlamos rezultātus.

Lai efektīvi pārvaldītu datu plūsmu, ne vienmēr ir nepieciešams algot vadītāju armiju, vispirms ir jāizstrādā stingras prasības un standarti datiem un jaizmanto atbilstoši rīki, lai tos automātiski validētu, saskaņotu un apstrādātu.



**Attēls 4.2-2 Lai nodrošinātu veselīgu uzņēmuma ekosistēmu, ir nepieciešams kvalitatīvs un savlaicīgs tās sistēmu nodrošinājums ar resursiem.**

Lai automatizētu datu validēšanas un saskaņošanas procesu (turpmākai automātiskai integrācijai), vispirms jāapraksta minimālās nepieciešamās datu prasības katrai konkrētai sistēmai. Šīs prasības nosaka:

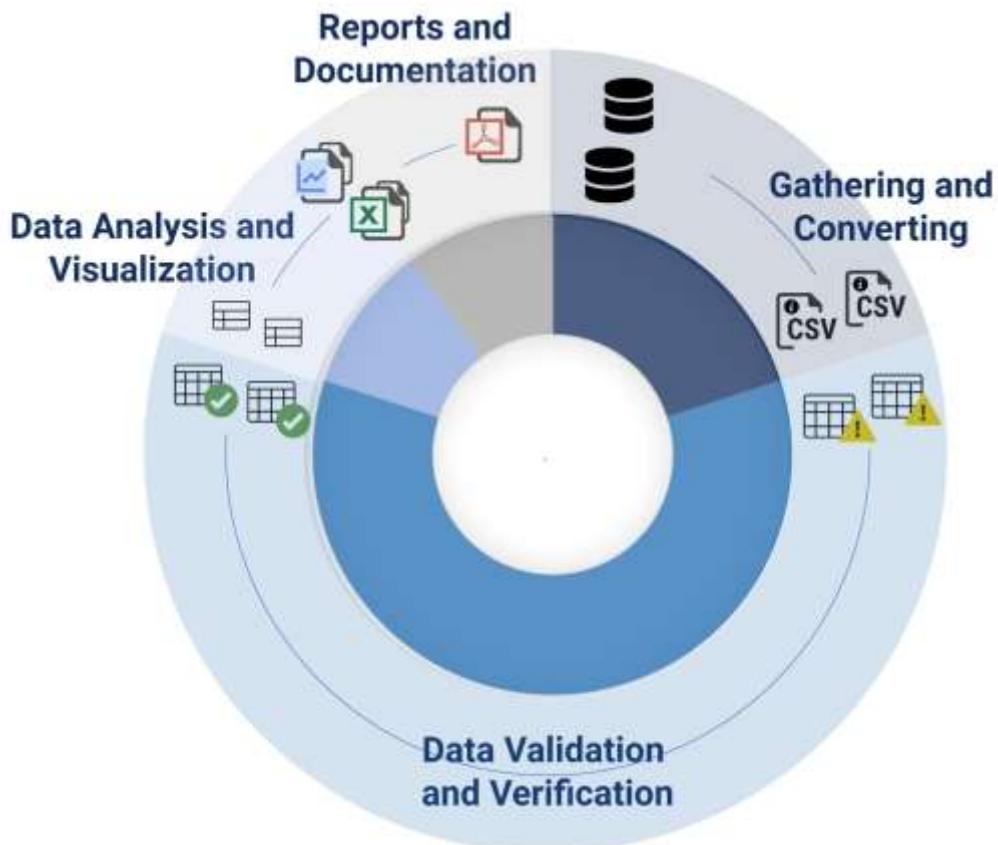
- Kas tieši jums ir nepieciešams?
- Kādā formā (struktūra, formāts)?
- Kādi atribūti ir obligāti?
- Kādas precizitātes un pilnīguma pielaides ir pieļaujamas?

Datu prasības apraksta saņemamās un apstrādājamās informācijas kvalitātes, struktūras un pilnīguma kritērijus. Piemēram, attiecībā uz tekstiem PDF -dokumentos ir svarīgi, lai tie būtu precīzi formatēti atbilstoši nozares standartiem ( 7.2-14. attēls - 7.2-16. attēls). Objektiem CAD -modeļos jābūt pareiziem atribūtiem (izmēriem, kodiem, saitēm ar klasifikatoriem) ( 7.3-9. att., 7.3-10. att.). Savukārt līgumu skenēšanai ir svarīgi skaidri datumi un iespēja automātiski iegūt summu un galvenos nosacījumus ( 4.1-7. attēls - 4.1-10. attēls).

Datu prasību formulēšana un automātiska to atbilstības pārbaude ir viens no laikietilpīgākajiem, bet būtiskākajiem soliem. Tas ir laikietilpīgākais uzņēmējdarbības procesu posms.

Kā minēts šīs grāmatas 3. daļā, no 50% līdz 90% biznesa izlūkošanas (BI) speciālistu laika tiek veltīts datu sagatavošanai, nevis analīzei ( 3.2-5. attēls). Šis process ietver datu vākšanu, pārbaudi, apstiprināšanu, saskaņošanu un strukturēšanu.

Saskaņā ar 2016. gadā veiktās aptaujas [95] rezultātiem datu zinātnieki dažādās plaša spektra jomās norādīja, ka lielāko daļu sava darba laika (aptuveni 80%) viņi pavada, darot to, kas viņiem vismazāk patīk ( 4.2-3. attēls): vācot esošās datu kopas un organizējot (apvienojot, strukturējot) tās. Tādējādi mazāk nekā 20% no viņu laika paliek radošiem uzdevumiem, piemēram, modeļu un likumsakarību meklēšanai, kas ļautu gūt jaunas atziņas un atklājumus.



Attēls 4.2-3 verifikācija un datu kvalitātes nodrošināšana ir visdārgākais, laikietilpīgākais un sarežģītākais posms datu sagatavošanā integrācijai citās sistēmās.

Veiksmīgai datu pārvaldībai būvniecības uzņēmumā ir nepieciešama visaptveroša pieeja, kas ietver uzdevumu parametrizēšanu, datu kvalitātes prasību formulēšanu un piemērotu rīku izmantošanu to automatizētai validēšanai.

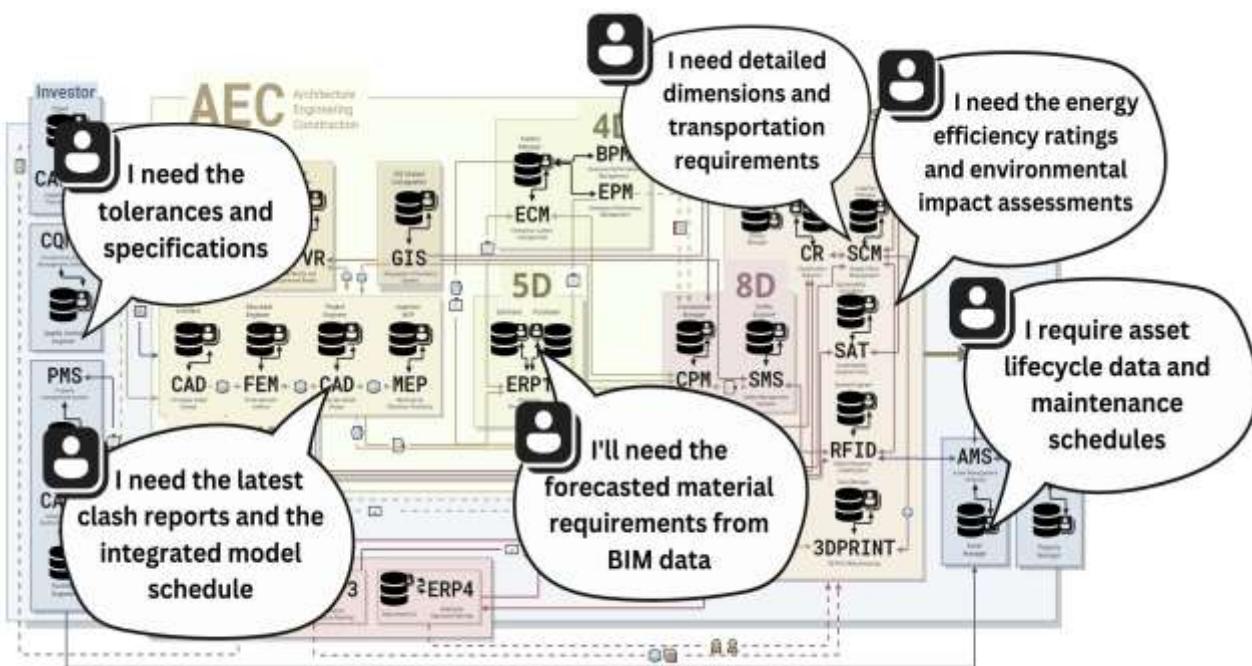
## Digitālā sadarbspēja sākas ar prasībām

Pieaugot digitālo sistēmu skaitam uzņēmumos, pieaug arī nepieciešamība pēc datu konsekvoņes starp tām. Par dažādām IT sistēmām atbildīgie vadītāji bieži vien nespēj sekot līdzi pieaugošajam informācijas apjomam un formātu daudzveidībai. Šādos apstākļos viņi ir spiesti lūgt speciālistus izveidot datus tādā formā, kas ir piemērota izmantošanai citās lietojumprogrammās un platformās.

Tas savukārt liek datu ģenerēšanā iesaistītajiem inženieriem un darbiniekim pielāgoties daudzām prasībām, bieži vien bez pārredzamības un skaistas izpratnes par to, kur un kā dati tiks izmantoti nākotnē. Standartizētu pieejumu trūkums informācijas apstrādei rada neefektivitāti un palielina izmaksas verifikācijas posmā, kas datu sarežģītības un nestandardizētā rakstura dēļ bieži ir manuāls.

Datu standartizācijas jautājums nav tikai ērtības vai automatizācijas jautājums. Tas ir tiešs finansiāls zaudējums. Saskaņā ar IBM 2016. gada ziņojumu gada zaudējumi, ko ASV rada slikta datu kvalitāte, ir 3,1 triljons ASV dolāru [96]. Turklat MIT un citu analītisko konsultāciju uzņēmumu pētījumi liecina, ka sliktas datu kvalitātes izmaksas var sasniegt pat 15-25% no uzņēmuma ieņēmumiem [97].

Šādos apstākļos klūst ļoti svarīgi, lai būtu skaidri definētas datu prasības un apraksti par to, kādi parametri, kādā formātā un ar kādu detalizācijas pakāpi būtu jāiekļauj izveidotajos objektos. Bez šo prasību formalizēšanas nav iespējams garantēt datu kvalitāti un savietojamību starp sistēmām un projekta posmiem ( 4.2-4. attēls).

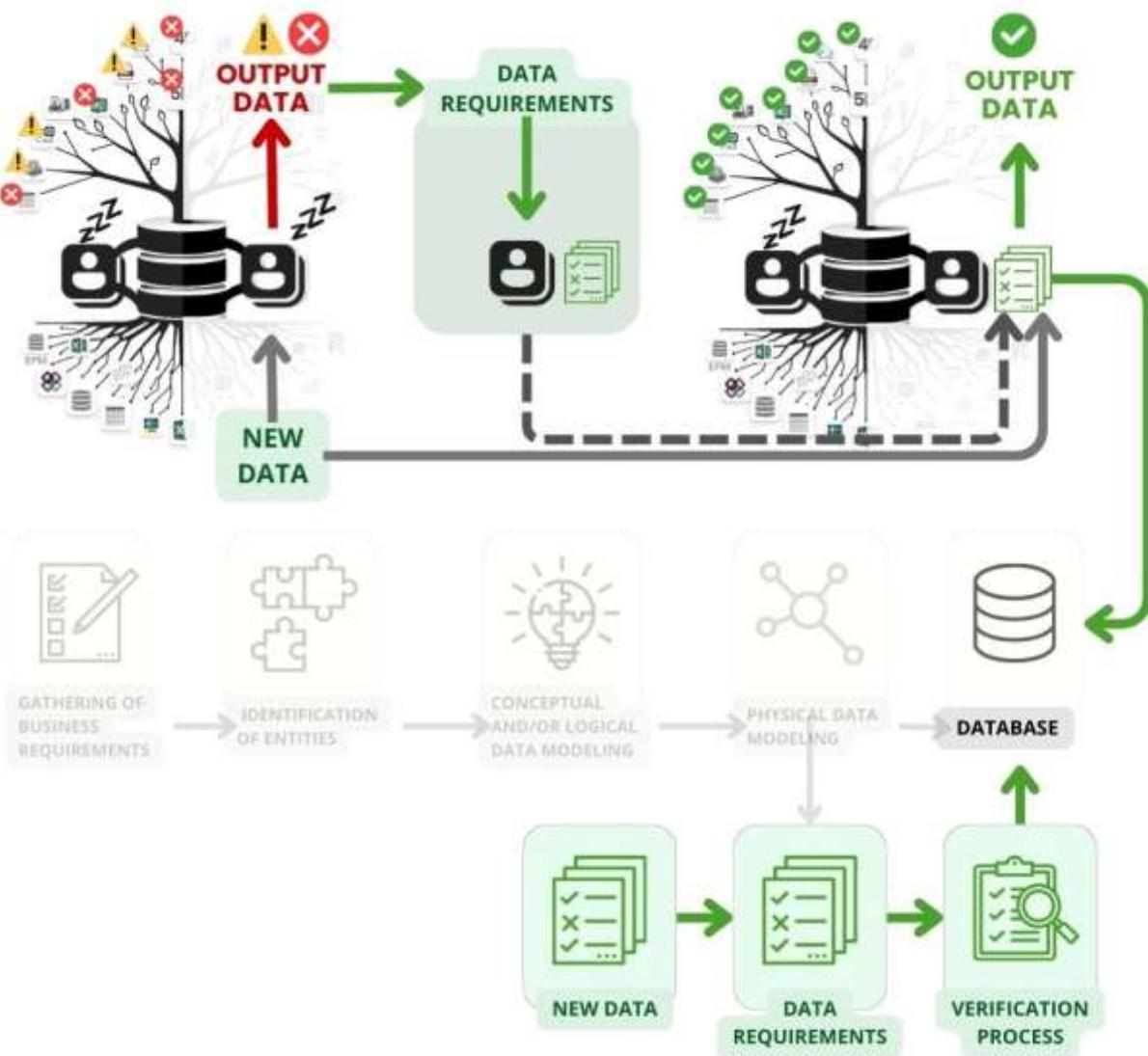


Attēls 4.2-4 Biznesa pamatā ir dažādu lomu mijiedarbība, un katrai no tām ir nepieciešami noteikti parametri un vērtības, kas ir kritiski svarīgi biznesa mērķu sasniegšanai.

Lai formulētu pareizas datu prasības, jums ir jāizprot uzņēmējdarbības procesi datu līmenī. Būvniecības projekti atšķiras pēc veida, mēroga un dalībnieku skaita, un katrai sistēmai - vai tā būtu

modelēšana (CAD (BIM)), plānošana (ERP 4D), izmaksu aprēķins (ERP 5D) vai loģistika (SCM) - ir nepieciešami savi unikāli ievades datu parametri (ievades entīžu elementi).

Atkarībā no šīm vajadzībām uzņēmumu vadītājiem ir vai nu jāizstrādā jaunas datu struktūras, lai atbilstu prasībām, vai jāpielāgo esošās tabulas un datubāzes. Izveidoto datu kvalitāte būs tieši atkarīga no tā, cik precīzi un pareizi formulētas prasības (4.2-5. attēls).



Attēls 4.2-5 Datu kvalitāte ir atkarīga no to prasību kvalitātes, kas tiek radītas konkrētiem datu izmantošanas gadījumiem.

Tā kā katrai sistēmai ir savas specifiskas datu prasības, pirms solis vispārējo prasību formulēšanā ir visu uzņēmējdarbības procesos iesaistīto elementu kategorizēšana. Tas nozīmē nepieciešamību objektus iedalīt klasēs un klašu grupās, kas atbilst konkrētām sistēmām vai lietojumprogrammas uzdevumiem. Katrai šādai grupai tiek izstrādātas atsevišķas prasības datu struktūrai, atribūtiem un

kvalitātei.

Tomēr praksē šīs pieejas īstenošana saskaras ar būtisku problēmu - trūkst vienotas valodas datu grupēšanai. Atšķirīgas klasifikācijas, dublēti identifikatori un nesaderīgi formāti rada to, ka katrs uzņēmums, katra programmatūra un pat katrs projekts veido savus, izolētus datu modeļus un klasses. Rezultātā veidojas digitālais "Bābeles tornis", kur informācijas pārsūtīšanai starp sistēmām ir nepieciešama vairākkārtēja konvertēšana uz pareizajiem datu modeļiem un klasēm, kas bieži tiek veikta manuāli. Šo barjeru var pārvarēt, tikai pārejot uz universāliem klasifikatoriem un standartizētiem prasību kopumiem.

## **Kopīga būvniecības valoda: klasifikatoru loma digitālajā pārveidē**

Pārbaudes un apstrādes procesu digitalizācijas un automatizācijas kontekstā īpaša loma ir klasifikācijas sistēmām - sava veida "digitālajām vārdnīcām", kas nodrošina vienotu objektu aprakstu un parametru noteikšanu. Klasifikatori veido "kopīgu valodu", kas ļauj datus sagrupēt pēc nozīmes un datus integrēt starp dažādām sistēmām, pārvaldības līmeņiem un projekta dzīves cikla posmiem.

Visjūtamākā klasifikatoru ietekme ir ēkas aprites cikla ekonomikā, kur vissvarīgākais aspekts ir ilgtermiņa ekspluatācijas izmaksu optimizācija. Pētījumi liecina, ka ekspluatācijas izmaksas veido līdz pat 80% no ēkas kopējām ekspluatācijas izmaksām, kas ir trīs reizes lielākas nekā sākotnējās būvniecības izmaksas (4.2-6. attēls) [98]. Tas nozīmē, ka lēmums par nākotnes izmaksām lielā mērā tiek pieņemts projektēšanas posmā

Tāpēc par izejas punktu datu prasību ġenerēšanai projektēšanas posmā būtu jākļūst darbības inženieru prasībām (CAFM, AMS, PMS, RPM) (1.2-4. attēls). Šīs sistēmas nebūtu jāuzskata par projekta noslēguma posmu, bet gan par neatņemamu visas projekta digitālās ekosistēmas sastāvdaļu no koncepcijas līdz demontāžai

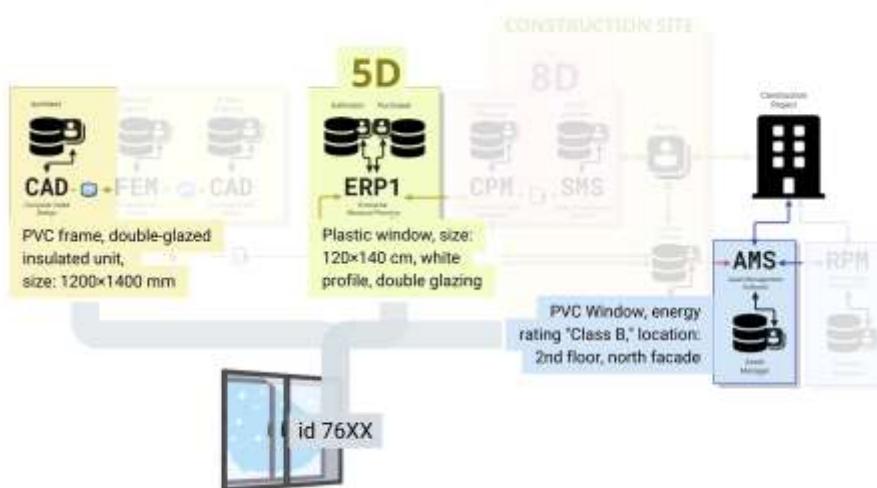
Mūsdienīgs klasifikators nav tikai grupēšanas kodu sistēma. Tas ir mehānisms savstarpējai sapratnei starp arhitektiem, inženieriem, tāmētājiem, loģistikas speciālistiem, tehniskās apkopes un IT sistēmām. Tāpat kā automašīnas autopilotam ir viennozīmīgi un precīzi jāatpazīst ceļa objekti, digitālajām būvniecības sistēmām un to lietotājiem ir viennozīmīgi jāinterpretē viens un tas pats projekta elements dažādām sistēmām, izmantojot elementu klasi.



Attēls 4.2-6 Ekspluatācijas un uzturēšanas izmaksas trīs reizes pārsniedz būvniecības izmaksas un veido 60-80% no visām ēkas aprites cikla izmaksām (saskaņā ar [99]).

Klasifikatoru izstrādes līmenis ir tieši saistīts ar uzņēmuma digitalizācijas pakāpi un tā digitālo briedumu. Organizācijas ar zemu digitālā brieduma līmeni saskaras ar sadrumstalotiem datiem, nesaderīgām informācijas sistēmām un līdz ar to nesaderīgiem un neefektīviem klasifikatoriem. Šādos uzņēmumos vienam un tam pašam elementam dažādās sistēmās bieži var būt dažādi grupēšanas identifikatori, kas būtiski kavē galīgo integrāciju un padara neiespējamu procesu automatizāciju.

Piemēram, viens un tas pats projekta logs CAD modelēšanas, tāmēšanas un uzturēšanas sistēmā var tikt apzīmēts atšķirīgi (4.2-7. attēls), jo dažādi procesa dalībnieki elementus uztver daudzdimENSIONĀLI. Aplēšu sastādītajam loga kategorijas elementā svarīgs ir apjoms un izmaksas, tehniskās apkopes dienestam - pieejamība un uzturējamība, arhitektam - estētiskie un funkcionālie raksturlielumi. Rezultātā vienam un tam pašam elementam var būt nepieciešami dažādi parametri.



Attēls 4.2-7 Ar nekonsekventu klasifikāciju starp sistēmām elements zaudēs daļu atribūtu informācijas katrā pārejas posmā uz citu sistēmu.

Tā kā ir grūti viennozīmīgi definēt būvelementu klasifikāciju, dažādu jomu speciālisti bieži vien vienam

un tam pašam elementam piešķir nesaderīgas klasses. Tā rezultātā tiek zaudēts vienots priekšstats par objektu, kas prasa turpmāku manuālu iejaukšanos, lai saskaņotu dažādās klasifikācijas sistēmas un panāktu konsekvenči starp dažādu speciālistu definētajiem tipiem un klasēm.

Šīs nekonsekences dēļ iepirkumu nodalas (ERP) saņemto operatīvo dokumentāciju, kad būvmateriāls tiek iegādāts no ražotāja, bieži vien nevar pareizi sasaistīt ar šī būvmateriāla klasifikāciju būvlaukumā (PMIS, SCM). Tā rezultātā kritiski svarīga informācija, visticamāk, netiks integrēta infrastruktūras un aktīvu pārvaldības sistēmās (CAFM, AMS), kas rada nopietnas problēmas nodošanas ekspluatācijā laikā, kā arī turpmākās apkopes (AMS, RPM) vai elementa nomaiņas laikā.

Uzņēmumos ar augstu digitālo briedumu klasifikatoriem ir nervu sistēmas loma, kas apvieno visas informācijas plūsmas. Viens un tas pats elements saņem unikālu identifikatoru, kas ļauj to bez izkroplojumiem vai zudumiem pārsūtīt starp CAD, ERP, AMS un CAFM sistēmām un to klasifikatoriem.

Lai izveidotu efektīvus klasifikatorus, ir jāsaprot, kā tiek izmantoti dati. Viens un tas pats inženieris dažādos projektos elementu var nosaukt un klasificēt atšķirīgi. Stabilu klasifikācijas sistēmu var izveidot tikai gadiem ilgi vācot lietošanas statistiku. Šajā darbā palīdz mašīnmācīšanās: algoritmi analizē tūkstošiem projektu (9.1-10. attēls), ar mašīnmācīšanās palīdzību nosakot iespējamās klasses un parametrus (10.1-6. attēls). Automātiskā klasifikācija ir īpaši vērtīga vidēs, kur manuāla klasifikācija nav iespējama datu apjoma dēļ. Automātiskās klasifikācijas sistēmas spēs atšķirt pamatkategorijas, pamatojoties uz minimāli aizpildītiem posteņu parametriem (sīkāka informācija devītajā un desmitajā grāmatas daļā).

Izstrādātās klasifikatoru sistēmas kļūst par katalizatoriem turpmākai digitalizācijai, radot pamatu:

- Automatizēta projekta izmaksu un termiņu aplēse.
- Potenciālo risku un konfliktu prognozēšanas analīze
- Iepirkuma procesu un loģistikas ķēžu optimizācija
- Ēku un struktūru digitālo dvīņu izveide
- Integrācija ar viedo pilsētu un lietu interneta sistēmām

Pārveides laiks ir ierobežots - attīstoties mašīnmācīšanās un datorredzes tehnoloģijām, automātiskās klasifikācijas problēma, kas gadu desmitiem ir bijusi neatrisināma, tuvākajos gados tiks atrisināta, un būvniecības un projektēšanas uzņēmumi, kas nespēs laikus pielāgoties, riskē atkārtot digitālo platformu izstumto taksometru parku likteni.

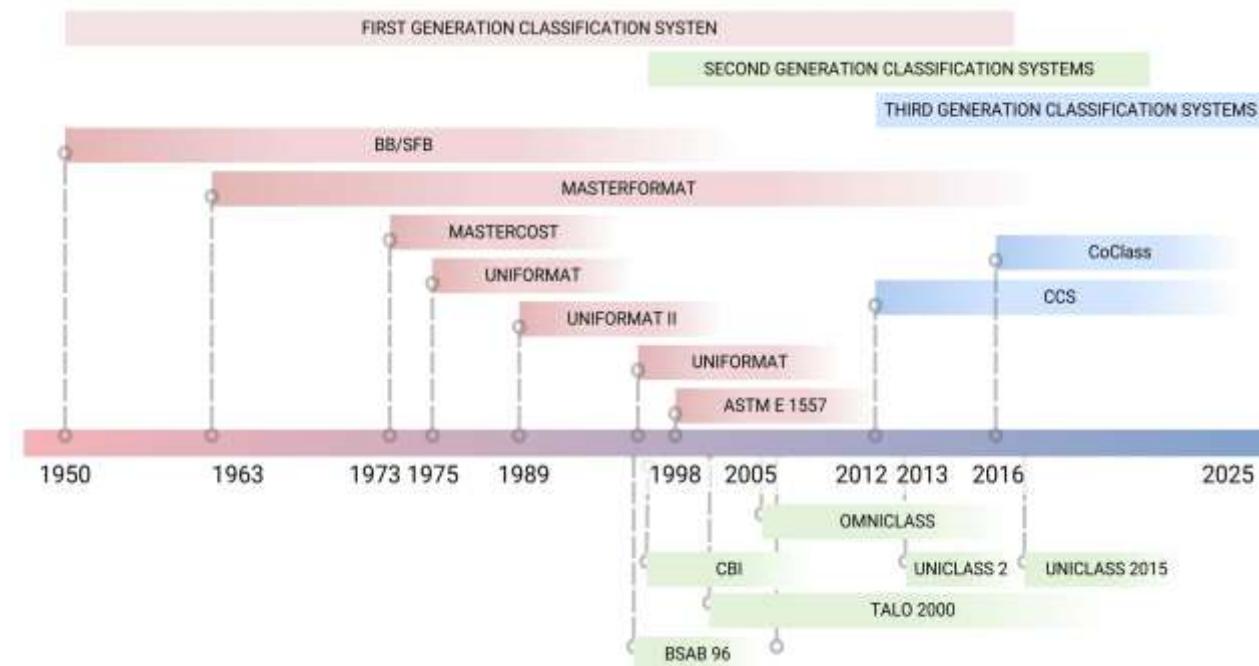
Izmaksu un laika grafiku aprēķinu automatizācija, kā arī lielie dati un mašīnmācīšanās tiks sīkāk aplūkota grāmatas piektajā un devītajā daļā. Risks, ka var atkārtoties taksometru autoparku liktenis un būvniecības nozares uberizācija, ir detalizēti aplūkots grāmatas desmitajā daļā.

Izprotot klasifikatoru būtisko lomu būvniecības nozares digitālās transformācijas procesā, ir nepieciešams pievērsties to attīstības vēsturei. Tieši vēsturiskais konteksts ļauj mums saprast, kā ir attīstījušās piejas klasifikatoriem un kādas tendences nosaka to pašreizējo stāvokli.

## Masterformat, OmniClass, Uniclass un CoClass: klasifikācijas sistēmu attīstība.

Vēsturiski būvniecības elementu un darbu klasifikatori ir attīstījušies trīs paaudzēs, no kurām katrā atspoguļo pieejamo tehnoloģiju līmeni un nozares aktuālās vajadzības konkrētā laika periodā ( 4.2-8. attēls):

- **Pirmā paaudze** (no 50. gadu sākuma līdz 80. gadu beigām) - papīra katalogi, hierarhiski klasifikatori, kurus izmantoja lokāli (piemēram, Masterformat, SfB).
- **Otrā paaudze** (no 90. gadu beigām līdz 2010. gadu vidum) ir elektroniskās tabulas un strukturētas datubāzes, kas ieviestas Excel un Access (ASTM E 1557, OmniClass, Uniclass 1997).
- **Trešā paaudze** (no 2010. gadiem līdz mūsdienām) - digitālie pakalpojumi un API - saskarnes, integrācija ar CAD (BIM), automatizācija (Uniclass 2015, CoClass).



Attēls 4.2-8 Trīs būvniecības nozares klasifikatoru paaudzes.

Pēdējās desmitgadēs ir samazinājusies klasifikatoru hierarhiskā sarežģītība ( 4.2-9. attēls): ja agrīnās sistēmas, piemēram, OmniClass, izmantoja līdz pat 7 ligzdošanas līmeņiem, lai aprakstītu 6887 klasses, tad mūsdienu risinājumi, piemēram, CoClass, aprobežojas ar 3 līmeņiem ar 750 klasēm. Tas atvieglo darbu ar datiem, vienlaikus saglabājot nepieciešamo granularitāti. Uniclass 2015, kas bieži tiek izmantots kā standarts Apvienotajā Karalistē, apvieno 7210 klasses tikai 4 līmeņos, padarot to ērtu CAD projektiem un publiskajiem iepirkumiem.

Classifier	Table / Objects	Number of classes	Nesting depth
OmniClass	Table 23 Products	6887	7 levels
Uniclass 2015	Pr — Products	7210	4 levels
CoClass, CCS	Components	750	3 levels

Attēls 4.2-9 Ar katru jaunu klasifikatoru paaudzi kategorizācijas sarežģītība strauji samazinās.

Dažādu valstu būvniecības tāmēšanas sistēmās atšķirīgo klasifikāciju dēļ pat tādu tipisku elementu kā betona pamatu sienai var aprakstīt ļoti atšķirīgi ( 4.2-10. attēls). Šīs atšķirības atspoguļo valstu būvniecības praksi, izmantotās mērišanas sistēmas, pieeju materiālu klasifikācijai un katrā valstī spēkā esošās normatīvās un tehniskās prasības.



Attēls 4.2-10 Dažādās valstīs projektos izmanto vienu un to pašu elementu, izmantojot dažādus aprakstus un klasifikācijas.

Tādu pašu elementu klasifikācijas dažādība sarežģī starptautisko sadarbību un padara izmaksu un darba apjoma salīdzināšanu starptautiskos projektos laikietilpīgu un dažkārt gandrīz neiespējamu. Pašlaik pasaules līmenī nav viena universāla klasifikatora - katrā valsts vai reģions izstrādā savas sistēmas, pamatojoties uz vietējām normām, valodu un uzņēmējdarbības kultūru:

- **CCS (Dānija):** Izmaksu klasifikācijas sistēma - sistēma izmaksu klasifikācijai visā iekārtas dzīves ciklā (projektēšana, būvniecība, ekspluatācija). Uzsvars tiek likts uz ekspluatācijas un uzturēšanas loģiku, bet ietver arī budžeta un resursu pārvaldību.
- **NS 3451 (Norvēģija):** kategorizē objektus pēc funkcijām, konstrukcijas elementiem un aprites cikla posmiem. Izmanto projektu vadībai, izmaksu aplēsēm un ilgtermiņa plānošanai.
- **MasterFormat (ASV):** sistēma būvniecības specifikāciju strukturēšanai sadaļās (piemēram, betona, elektrības, apdares). Koncentrējas uz disciplīnām un darbu veidiem, nevis funkcionāliem elementiem (atšķirībā no UniFormat).
- **Uniclass 2 (Apvienotā Karaliste):** viens no detalizētākajiem klasifikatoriem, ko izmanto publiskā

iepirkuma un BIM projektos. Apvieno datus par objektiem, darbiem, materiāliem un telpām vienotā sistēmā.

- **OmniClass:** starptautisks standarts (izstrādājis CSI ASV) objektu informācijas pārvaldībai no komponentu bibliotēkām līdz elektroniskajām specifikācijām. Piemērots ilgtermiņa datu uzglabāšanai, saderīgs ar CAD (BIM) un citiem digitālajiem rīkiem.
- **COBie:** Būvniecības un ekspluatācijas informācijas apmaiņa ir starptautisks standarts datu apmaiņai starp projektēšanas, būvniecības un ekspluatācijas posmiem. Iekļauts BS 1192-4:2014 kā daļa no koncepcijas "BIM - lietošanai gatavs modelis". Koncentrējas uz informācijas nodošanu (piemēram, iekārtu specifikācijas, garantijas, kontakti ar darbuzņēmējiem).

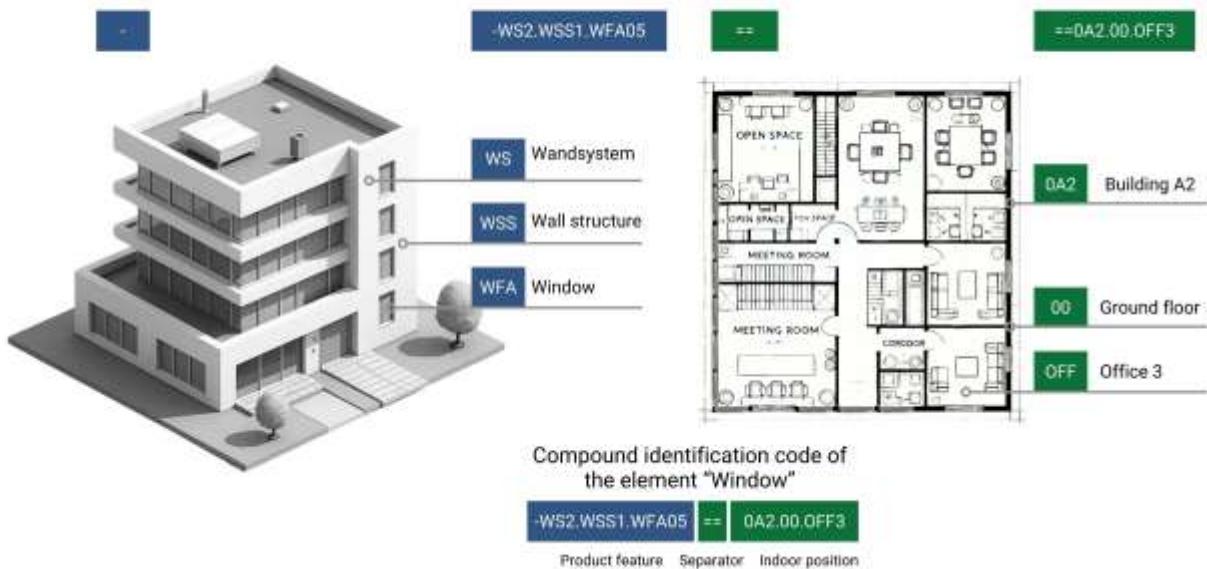
Būvniecības nozares globalizācija, visticamāk, veicinās pakāpenisku būvelementu klasifikācijas sistēmu unifikāciju, kas ievērojami samazinās atkarību no vietējiem valsts standartiem. Šis process var attīstīties līdzīgi interneta komunikāciju attīstībai, kad universālie datu pārsūtīšanas protokoli galu galā izspieda atšķirīgos vietējos formātus, nodrošinot globālu sistēmu savietojamību.

Alternatīvs attīstības ceļš var būt tieša pāreja uz automātiskās klasifikācijas sistēmām, kuru pamatā ir mašīnmācīšanās tehnoloģijas. Šīm tehnoloģijām, kas pašlaik tiek attīstītas galvenokārt autonomā transporta jomā, ir ievērojams potenciāls, lai tās piemērotu lielām CAD projektēšanas datu kopām (10.1-6. attēls).

Šobrīd situācija neaprobežojas tikai ar klasifikatoru grupēšanu valstī. Daudzo īpatnību dēļ, kas nav ķemtas vērā valsts līmenī, katram uzņēmumam ir jāunificē un jāstandartizē elementu un resursu kategorijas, ar kurām tas strādā.

Parasti šis process parasti sākas ar nelielām - vietējām objektu tabulām vai iekšējām marķēšanas sistēmām. Tomēr stratēģiskais mērķis ir pāriet uz vienotu valodu visu elementu aprakstīšanai, kas būtu saprotama ne tikai uzņēmumā, bet arī ārpus tā - ideālā gadījumā saskaņota ar starptautiskiem vai nozares klasifikatoriem ( 4.2-8. attēls). Šāda pieeja atvieglo integrāciju ar ārējiem partneriem, digitālajām sistēmām un veicina vienotu kompleksu procesu veidošanos objektu dzīves ciklā.

Pirms pāriet uz automatizāciju un mērogojamām IT sistēmām, ir jāizmanto valsts līmeņa klasifikatori vai jāizveido sava loģiska un nepārprotama elementu identifikācijas struktūra. Katrs objekts - vai tas būtu logs ( 4.2-11. attēls), durvis vai inženiersistēma - ir jāapraksta tā, lai to varētu nepārprotami atpazīt jebkurā uzņēmuma digitālajā sistēmā. Tas ir ļoti svarīgi, pārejot no plakaniem rasējumiem uz digitālajiem modeļiem, kas aptver gan projektēšanas posmu, gan ēkas ekspluatāciju.



**Attēls 4.2-11 Saliktā loga ēkas elementa identifikatora piemērs, pamatojoties uz klasifikāciju un atrašanās vietu ēkā.**

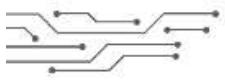
Viens no iekšējo klasifikatoru piemēriem varētu būt saliktā identifikācijas koda izveide ( 4.2-11. attēls). Šāds kods apvieno vairākus informācijas līmeņus: elementa funkcionālo funkciju (piemēram, "logs sienā"), tā tipu un precīzu telpisko norādi - ēka A2, 0. stāvs, 3. telpa. Šāda daudzīmēju struktūra ļauj izveidot vienotu navigācijas sistēmu digitālajos modeļos un dokumentācijā, jo īpaši datu pārbaudes un transformācijas posmos, kur nepieciešama viennozīmīga elementu grupēšana. Viennozīmīga elementu atpazīšana nodrošina konsekvenci starp nodaļām un samazina dublēšanās, kļūdu un informācijas zuduma risku.

Labi izstrādāts klasifikators nav tikai tehnisks dokuments, tas ir uzņēmuma digitālās ekosistēmas pamats:

- nodrošina datu savietojamību starp sistēmām;
- samazina informācijas meklēšanas un apstrādes izmaksas;
- palielina pārredzamību un pārvaldāmību;
- rada pamatu mērogošanai un automatizācijai.

Standartizēts objektu apraksts, izmantojot valsts klasifikatorus vai patentētus saliktos identifikācijas kodus, kļūst par pamatu konsekventiem datiem, uzticamai informācijas apmaiņai un turpmākai viedo pakalpojumu ieviešanai - no automatizēta iepirkuma līdz digitālajiem dvīņiem.

Pēc daudzformāta datu strukturēšanas posma pabeigšanas un klasifikatora, kas tiks izmantots elementu atpazīšanai un grupēšanai, izvēles nākamais solis ir pareizi modelēt datus. Šis process ietver galveno parametru noteikšanu, loģiskas datu struktūras izveidi un attiecību starp elementiem



aprakstīšanu.

## NODAĻA 4.3.

### DATU MODELĒŠANA UN IZCILĪBAS CENTRS

#### Datu modelēšana: konceptuālais, loģiskais un fiziskais modelis

Efektīva datu pārvaldība (kurus mēs iepriekš strukturējām un klasificējām) nav iespējama bez pārdomātas glabāšanas un apstrādes struktūras. Lai nodrošinātu piekļuvi informācijai un tās konsekvenci glabāšanas un apstrādes posmā, uzņēmumi izmanto datu modelēšanu - metodoloģiju, kas ļauj izstrādāt tabulas, datubāzes un saites starp tām atbilstoši uzņēmējdarbības prasībām.

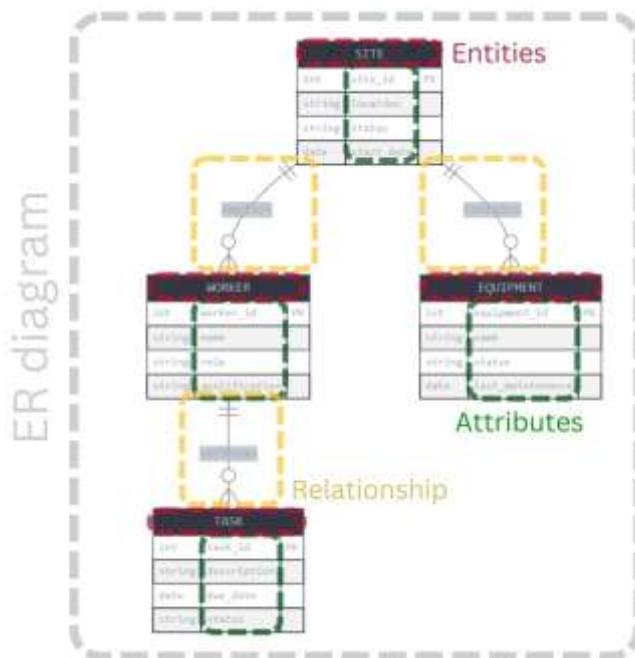
Datu modelēšana ir pamats, uz kura tiek veidota jebkura digitālā ekosistēma. Bez sistēmu, prasību un datu modelēšanas apraksta inženieri un speciālisti, kas veido datus, nezina un nesaprot, kur viņu radītie dati tiks izmantoti.

Līdzīgi kā būvējot ēku, kur nevar sākt kieģeļu likšanu bez plāna, arī datu noliktavas sistēmas izveidei nepieciešama skaidra izpratne par to, kādi dati tiks izmantoti, kā tie tiks saistīti un kas ar tiem strādās. Ja nav procesu un prasību apraksta, inženieri un speciālisti, kas veido datus, zaudē ieskatu, kur un kā dati tiks izmantoti nākotnē.

Datu modelis kalpo kā tilts starp uzņēmējdarbību un IT. Tas ļauj formalizēt prasības, strukturēt informāciju un atvieglot saziņu starp ieinteresētajām personām. Šajā ziņā datu modelēšana ir līdzīga arhitekta darbam, kurš saskaņā ar pasūtītāja plānu izstrādā ēkas plānu un pēc tam nodod to būvniekiem - datubāzu administratoriem un izstrādātājiem - īstenošanai (datubāzes izveidei).

Tādējādi ikvienam būvniecības uzņēmumam papildus elementu un resursu strukturēšanai un kategorizēšanai (4.2-11. attēls) ir jāapgūst datu bāzu (tabulu) "būvēšanas" māksla un jāiemācās veidot saiknes starp tām, it kā savienojot kieģeļus uzticamā un spēcīgā zināšanu sienā no uzņēmuma datiem. Galvenie datu modelēšanas jēdzieni (4.3-1. attēls) ir šādi:

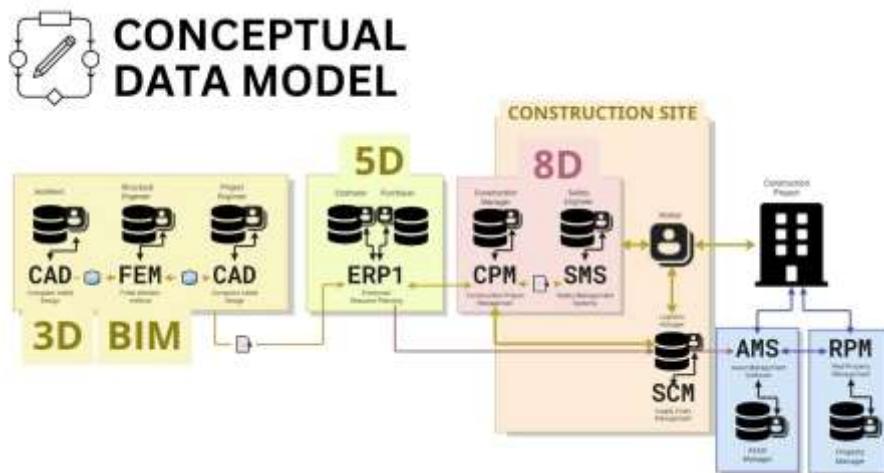
- **Entitātes** ir objekti, par kuriem jāvāc dati. Sākotnējā projektēšanas posmā vienība var būt atsevišķs elements (piemēram, "durvis"), bet aprēķinu modelī tā var būt kategorizētu elementu grupa (piemēram, "iekšdurvis").
- **Atribūti** ir vienību raksturlielumi, kas apraksta svarīgas detaļas: izmērus, īpašības, montāžas izmaksas, loģistikas un citus parametrus.
- **Attiecības (saites)** - parāda, kā vienības savstarpēji mijiedarbojas. Tās var būt viena veida: "viens ar vienu", "daudzi ar vienu", "daudzi ar daudziem".
- **ER diagrammas** (Entity-Relationship diagrammas) ir vizuālas diagrammas, kurās parādītas entitātes, atribūti un attiecības starp tām. ER diagrammas var būt konceptuālas, loģiskas un fiziskas - katra no tām atspoguļo atšķirīgu detalizācijas līmeni.



Attēls 4.3-1 konceptuālas datubāzes struktūras ER diagramma ar vienībām, atribūtiem un attiecībām.

Datu projektēšanas un attiecību starp tiem definēšanas procesu tradicionāli iedala trīs galvenajos modeļos. Katrs no tiem pilda noteiktas funkcijas, un tie atšķiras pēc detalizācijas līmeņa un abstrakcijas pakāpes datu struktūras attēlojumā:

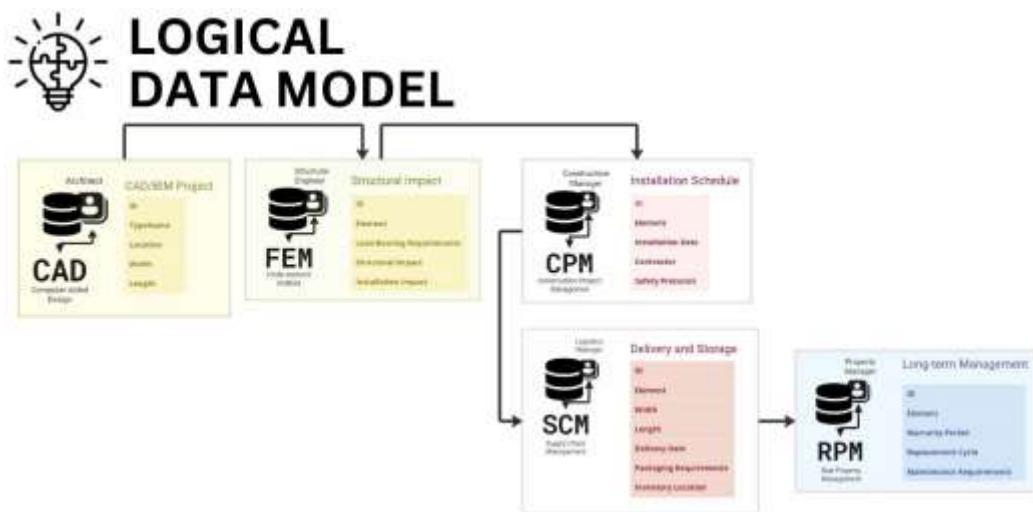
- **Konceptuālais datu modelis:** šis modelis apraksta galvenās vienības un to attiecības, neiedzīlinoties detaļās. To parasti izmanto plānošanas sākumposmā. Šajā posmā mēs varam iekarīt no datubāzēm un sistēmām, lai parādītu attiecības starp dažādām nodaļām un speciālistiem.



Attēls 4.3-2 Konceptuālā diagramma apraksta sistēmas saturu: augsta līmeņa attiecību attēlojums bez tehniskām detaļām.

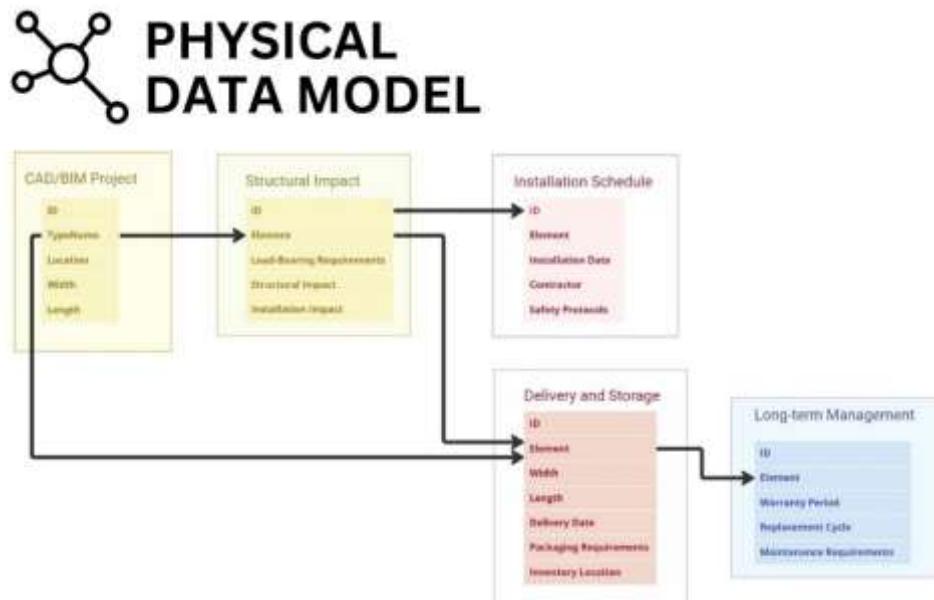
- **Loģiskais datu modelis:** Balstoties uz konceptuālo modeli, loģiskais datu modelis ietver

detalizētus būtnu, atribūtu, atslēgu un attiecību aprakstus, uzņēmējdarbības informācijas un noteikumu kartēšanu.



Attēlā 4.3-3 Logiskais datu modelis ir detalizēti norādīti datu tipi, sakarības un atslēgas, bet bez sistēmas implementācijas.

- Fiziskais datu modelis:** Šajā modeļā ir aprakstītas nepieciešamās struktūras datu bāzes īstenošanai, tostarp tabulas, kolonas un attiecības. Tajā galvenā uzmanība pievērsta datubāzes veiktspējai, indeksēšanas stratēģijām un fiziskajai glabāšanai, lai optimizētu datubāzu fizisko izvietošanu.

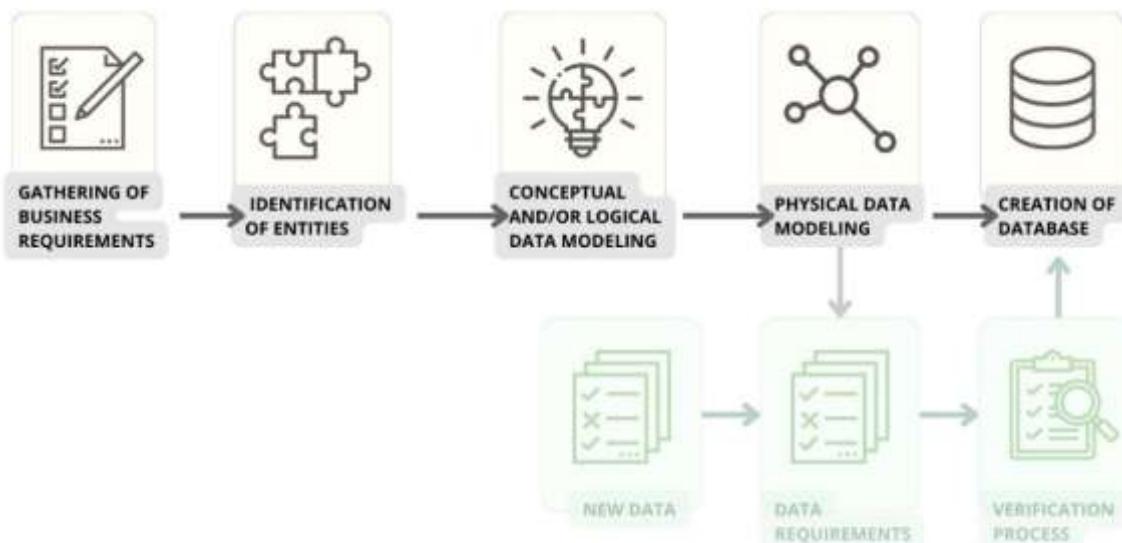


Attēls 4.3-4 Fiziskais datu modelis nosaka, kā tiks īstenota sistēma, tostarp tabulas un konkrētas datu bāzes detaļas.

Veidojot datubāzes un projektējot tabulārās attiecības, izpratnei par abstrakcijas līmeņiem ir būtiska nozīme efektīvas sistēmas arhitektūras izveidē.

Efektīva datu modelēšanas metodoloģija ļauj apvienot uzņēmējdarbības mērķus ar tehnisko īstenošanu, padarot visu procesu kēdi pārredzamāku un vieglāk vadāmu. Datu modelēšana nav vienreizējs uzdevums, bet process, kas ietver secīgus soļus (4.3-5. attēls):

- **Uzņēmējdarbības prasību apkopošana:** tiek definēti galvenie uzdevumi, mērķi un informācijas plūsmas. Šajā posmā notiek aktīva mijiedarbība ar ekspertiem un lietotājiem.
- **Entitāšu identifikācija:** tiek izcelti galvenie objekti, kategorijas un datu tipi, kurus ir svarīgi įņemt vērā nākotnes sistēmā.
- **Konceptuālā un loģiskā modeļa izveide:** vispirms tiek attēlotas galvenās vienības un to attiecības, pēc tam atrībūti, noteikumi un detalizēta struktūra.
- **Fiziskā modelēšana:** tiek izstrādāta modeļa tehniskā implementācija: tabulas, lauki, attiecības, ierobežojumi, indeksi.
- **Datubāzes izveide:** pēdējais solis ir fiziskā modeļa ieviešana izvēlētajā DBVS, testēšana un sagatavošana darbam.



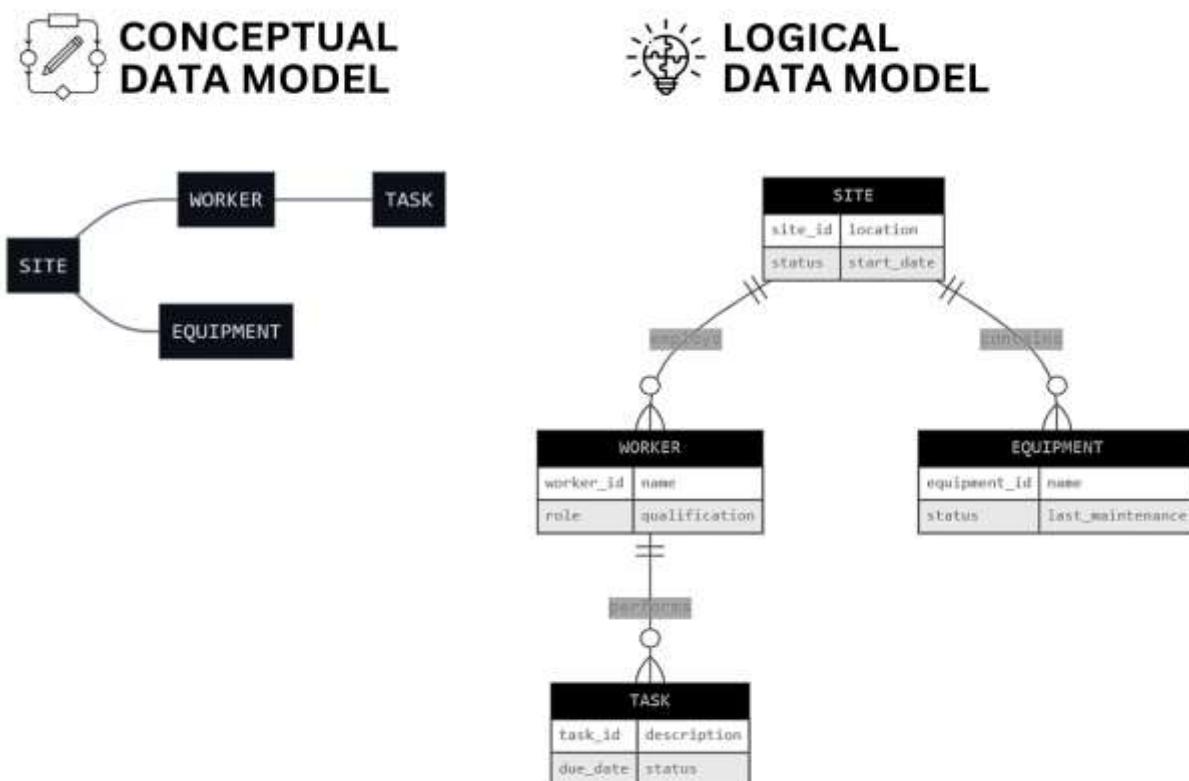
Attēls 4.3-5 Datu bāzu un datu pārvaldības sistēmu izveide biznesa procesiem sākas ar prasību ģenerēšanu un datu modelēšanu.

Pareizi izstrādāti datu modelēšanas procesi ļauj nodrošināt pārredzamas informācijas plūsmas, kas ir īpaši svarīgi sarežģītos projektos, piemēram, būvniecības projektu vai būvlaukumu pārvaldībā. Aplūkosim, kā pāreja no konceptuālā modeļa uz loģisko modeli un pēc tam uz fizisko modeli var palīdzēt rationalizēt procesus.

## Praktiska datu modelēšana būvniecības kontekstā

Kā datu modelēšanas piemēru ņemsim būvlaukuma vadības uzdevumu un konvertēsim brigadiera prasības strukturētā loģiskajā modelī. Pamatojoties uz būvlaukuma pārvaldības pamatvajadzībām, mēs definējam galvenās vienības: būvlaukumam (SITE), darbiniekiem (WORKER), aprīkojumam (EQUIPMENT), uzdevumiem (TASK) un aprīkojuma izmantošanai (EQUIPMENT\_USAGE). Katrai vienībai ir atribūtu kopums, kas atspoguļo svarīgas īpašības. Piemēram, TASK gadījumā tas varētu būt uzdevuma apraksts, izpildes termiņš, statuss, prioritāte; WORKER gadījumā tas varētu būt nosaukums, loma objektā, pašreizējā nodarbinātība utt.

Loģiskais modelis nosaka attiecības starp šīm vienībām, parādot, kā tās savstarpēji mijiedarbojas reālajos darba procesos (4.3-6. attēls). Piemēram, saikne starp vietu un darbiniekiem norāda, ka vienā vietā var strādāt daudzi darbinieki, savukārt saikne starp darbiniekiem un uzdevumiem atspoguļo, ka viens darbinieks var veikt vairākus uzdevumus.



Attēls 4.3-6 Konceptuālais un loģiskais datu modelis, kas izveidots pēc meistara prasībām būvlaukuma procesu aprakstīšanai.

Pārejot uz fizisko modeli, tiek pievienotas tehniskās īstenošanas detaļas: specifiski datu tipi (VARCHAR, INT, DATE), primārie un ārējie atslēgas attiecībām starp tabulām, kā arī indeksi, lai optimizētu datubāzes veikspēju (4.3-7. attēls).

Piemēram, statusiem jādefinē konkrēti tipi ar iespējamām vērtībām un jāpievieno indeksi galvenajiem laukiem, piemēram, status un worker\_id, lai uzlabotu meklēšanas veikspēju. Tas pārvērš sistēmas loģisko aprakstu konkrētā datubāzes īstenošanas plānā, kas ir gatavs izveidei un ieviešanai.

# PHYSICAL DATA MODEL



Attēls 4.3-7 Fiziskais datu modelis apraksta būvlaukuma vienības, izmantojot minimālos nepieciešamos parametrus.

Fiziskais modelis bieži atšķiras no loģiskā modeļa. Vidēji modelēšanas laiks tiek sadalīts šādi: aptuveni 50% tiek tērēti konceptuālajam modelim (prasību apkopošana, procesu apspriešana, vienību identificēšana), 10% - loģiskajam modelim (atribūtu un attiecību precizēšana) un 40% - fiziskajam modelim (ieviešana, testēšana, pielāgošana DBVS).

Šāds līdzsvars ir saistīts ar to, ka konceptuālajā posmā tiek likts pamats datu struktūrai, bet loģiskajā modelī tikai precizētas attiecības un atribūti. Fiziskais modelis prasa visvairāk resursu, jo tieši šajā posmā dati tiek ieviesti konkrētās platformās un rīkos.

## Datu bāzes izveide, izmantojot LLM

Kad ir izveidots datu modelis un ar parametru palīdzību aprakstītas vienības, mēs esam gatavi izveidot datu bāzes - krātuves, kurās uzglabāsim informāciju, kas iegūta pēc strukturēšanas posma par konkrētiem procesiem.

Mēģināsim izveidot vienkāršas, bet funkcionālas datubāzes piemēru ar minimālu koda apjomu, izmantojot SQLite, izmantojot Python programmēšanas valodas piemēru. Relāciju datubāzes tika detalizēti aplūkotas nodaļā "Strukturētās relāciju datubāzes un SQL vaicājumu valoda".

SQLite ir viegla, atklātā pirmkoda, iebūvējama, starplīdzīga datubāze, kas atbalsta SQL pamatlīdzību. SQLite tiek plaši izmantots CAD (BIM), ERP, EPM un ir gandrīz katrā viedtālrunī.

Palūgsim LLM modelim izveidot mums nelielu SQLite datubāzi, pamatojoties uz iepriekš izveidoto fizisko modeli ( 4.3-4. attēls) FEM (strukturālā ietekme) un CPM (montāžas grafiks) sistēmām.

- ✍ Nosūtiet teksta pieprasījumu LLM tērzēšanai (ChatGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkuru citu):

Izveidojiet SQLite datubāzes sistēmām "Konstruktīvā ietekme" un "Uzstādīšanas grafiks", izmantojot pielikumā pievienoto attēlu (ielīmējiet shēmu tērzētavā kā attēlu ( 4.3-4. attēls)»

2 LLM atbilde :



```
1 import sqlite3
2
3 # Connect to SQLite database (or create if it doesn't exist)
4 conn = sqlite3.connect('construction.db')
5
6 # Create a cursor object using the cursor() method
7 cursor = conn.cursor()
8
9 # Create table - Structural Impact
10 cursor.execute('''CREATE TABLE IF NOT EXISTS StructuralImpact (
11     ID INTEGER PRIMARY KEY,
12     Element TEXT,
13     LoadbearingRequirements TEXT,
14     StructuralImpact TEXT,
15     InstallationImpact TEXT
16 )''' )
17
18 # Create table - Installation Schedule
19 cursor.execute('''CREATE TABLE IF NOT EXISTS InstallationSchedule (
20     ID INTEGER PRIMARY KEY,
21     Element TEXT,
22     InstallationDate DATE,
23     Contractor TEXT,
24     SafetyProtocols TEXT
25 )''' )
26
27 # Commit the changes
28 conn.commit()
29
30 # Close the connection
31 conn.close()
32
33 print("Database and tables created successfully.")
```

Run in IDE

Attēls 4.3-8 Izgenerētais kods no augšupielādētā attēla izveido datu bāzes un tabulas, kurās tagad varam ielādēt datus.

Šo datubāzes kodu var palaist Python vidē ar SQLite atbalstu, kas parasti ir iekļauts standarta Python vidē. To var palaist jebkurā vietējā Python datorā, kas spēj palaist skriptus un izveidot failus. Pēc palaišanas un izpildes šis kods cietajā diskā izveidos failu - SQLite datubāzi ar nosaukumu construction.db (līdzīgi kā Excel failā ar darblapām), kurā būs tabulas "Strukturālā ietekme" un "Uzstādīšanas grafiks".

Kad šajās tabulās ir izveidota jauna datubāze **construction.db**, datus šajās tabulās var pievienot, izmantojot SQL\_ieprasījumus vai importu, kas vēlāk ļaus sākt automātisku datu apstrādi. Datus SQLite datubāzē var importēt no CSV failiem, Excel izklājlapām vai eksportēt no citām datubāzēm un krātuvēm, izmantojot API.

Lai izveidotu ilgtspējīgu datu modelēšanu un efektīvus datu bāzu pārvaldības procesus, uzņēmumam ir nepieciešama skaidri definēta stratēģija, kā arī koordinācija starp tehniskajām un biznesa komandām. Nēmot vērā atšķirīgus projektus un vairākus datu avotus, bieži vien ir grūti nodrošināt konsekvenči, standartizāciju un kvalitātes kontroli visos līmeņos. Viens no svarīgākajiem risinājumiem var būt īpaša datu modelēšanas izcilības centra (CoE) izveide uzņēmumā.

## Datu modelēšanas izcilības centrs (CoE)

Tā kā dati kļūst par vienu no galvenajiem stratēģiskajiem aktīviem, uzņēmumiem ir jādara vairāk nekā tikai pareizi vākt un uzglabāt informāciju - ir svarīgi iemācīties sistemātiski pārvaldīt datus. Klasifikācijas un datu modelēšanas izcilības centrs (CoE) ir struktūrvienība, kas nodrošina visu datu apstrādes konsekvenči, kvalitāti un efektivitāti organizācijā.

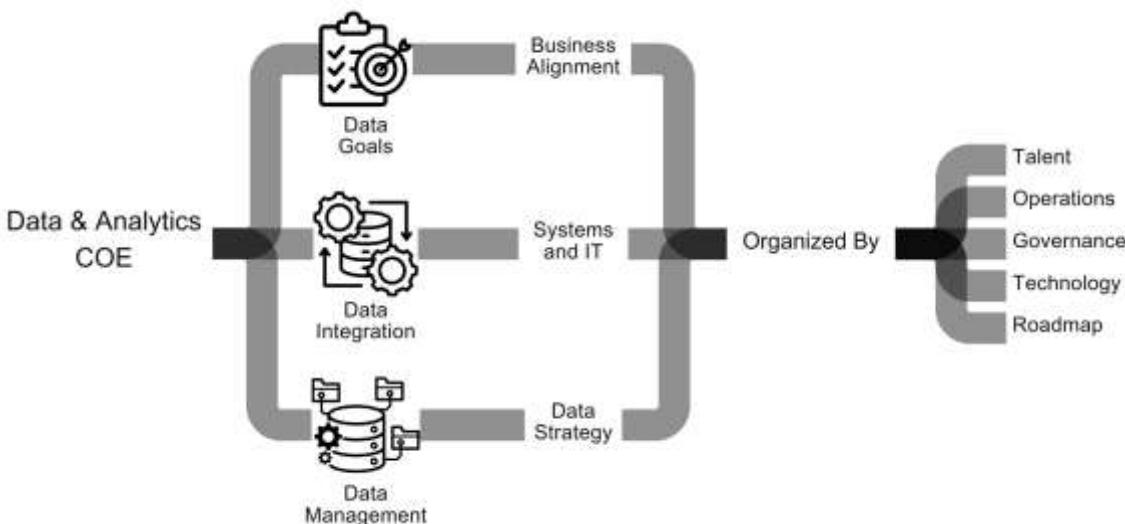
Izcilības centrs (CoE) ir ekspertu atbalsta kodols un uzņēmuma digitālās transformācijas metodoloģiskais pamats. Tas veido uz datiem balstītu kultūru un ļauj organizācijām veidot procesus, kas pieņem lēmumus, pamatojoties uz strukturētiem, validētiem un reprezentatīviem datiem, nevis intuīciju vai vietējo informāciju.

Datu izcilības centru parasti veido starpfunkcionālas komandas, kas strādā pēc "divu picu" principa. Šis Džefa Bezosa ierosinātais princips nozīmē, ka komandas lielumam jābūt tādam, lai to varētu pabarot ar divām picām, t. i., ne vairāk par 6-10 cilvēkiem. Šī pieja paīdz Izvairīties no pārmērīgas birokrātijas un palielina darba elastību. CoE komandā jāiekļauj darbinieki ar dažādām tehniskām prasmēm, sākot no datu analītikiem un mašīnmācīšanās un beidzot ar zināšanām konkrētās uzņēmējdarbības jomās. Ar savām padziļinātajām tehniskajām zināšanām datu inženieriem ne tikai jāoptimizē procesi un jāmodelē dati, bet arī jāatbalsta kolēģi, samazinot rutīnas uzdevumu veikšanas laiku (4.3-9. attēls).

Tāpat kā dabā ekosistēmas elastību nodrošina bioloģiskā daudzveidība, arī digitālajā pasaule elastību un pielāgošanās spēju nodrošina daudzveidīgas pieejas datu apstrādei. Tomēr šīs daudzveidības pamatā jābūt kopīgiem noteikumiem un koncepcijām.

Izcilības centru (CoE) var salīdzināt ar meža ekosistēmas "klimatiskajiem apstākļiem", kas nosaka, kuri

datu veidi plaukst un kuri tiek automātiski noraidīti. Radot labvēlīgu "klimatu" kvalitatīviem datiem, CoE veicina dabisku paraugprakses un metodoloģiju atlasi, kas pēc tam kļūst par organizācijas standartiem.



**Attēls 4.3-9** Datu un analītikas izcilības centrs (CoE) apvieno zināšanas par galvenajiem datu pārvaldības, integrācijas un stratēģijas aspektiem.

Lai paātrinātu integrācijas ciklus un sasniegtu labākus rezultātus, Eiropas Ekonomikas komitejai jānodrošina tās locekļiem pietiekama autonomija lēmumu pieņemšanā. Tas ir īpaši svarīgi dinamiskā vidē, kur izmēģinājumi un kļūdas, pastāvīga atgriezeniskā saite un biežas publikācijas var dot ievērojamu labumu. Tomēr šī autonomija ir efektīva tikai tad, ja ir skaidra saziņa un atbalsts no augstākās vadības puses. Bez stratēģiska redzējuma un augstākā līmeņa koordinācijas pat viskompetentākā komanda var saskarties ar šķēršļiem savu iniciatīvu īstenošanā.

Tieši COE vai uzņēmuma augstākā vadība ir atbildīga par to, lai datu modelēšanas pieeja neaprobežotos tikai ar vienu vai diviem projektiem, bet tiktu integrēta kopējā informācijas pārvaldības un uzņēmējdarbības procesu vadības sistēmā.

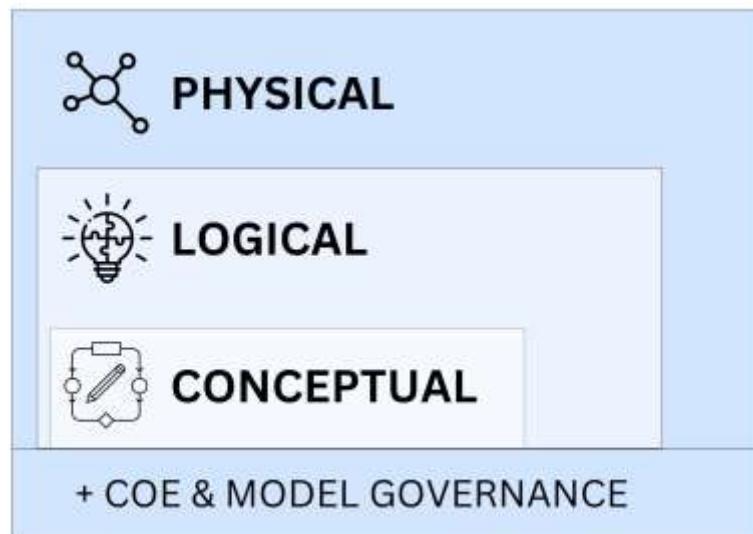
Papildus uzdevumiem, kas saistīti ar datu modelēšanu un datu pārvaldību, Ekspertīzes centrs (EPC) ir atbildīgs arī par vienotu standartu un pieeju izstrādi datu infrastruktūras izvietošanai un darbībai. Turklāt tas veicina nepārtrauktu uzlabojumu kultūru, procesu optimizāciju un datu efektīvu izmantošanu organizācijā (4.3-10. attēls).

Sistemātisko pieeju datu un modeļu pārvaldībai CoE var aptuveni iedalīt vairākos galvenajos blokos:

- **Procesu standartizācija un modeļu dzīves cikla pārvaldība:** CoE izstrādā un ievieš metodoloģijas, lai unificētu datu modeļu izveidi un pārvaldību. Tas ietver: strukturālo veidņu, kvalitātes kontroles metožu un versiju kontroles sistēmu izveidi, lai nodrošinātu datu nepārtrauktību visos darba posmos.
- **Lomu pārvaldība un atbildības piešķiršana:** CoE nosaka galvenās lomas datu modelēšanas procesā. Katram projekta dalībniekam ir skaidri noteiktas lomas un atbildības jomas, kas

atvieglo komandas darbu un samazina datu neatbilstību risku.

- **Kvalitātes kontrole un revīzija:** efektīvi būvniecības datu pārvaldībai nepieciešama nepārtraukta to kvalitātes uzraudzība. Tieki ieviesti automatizēti mehānismi datu pārbaudei, kļūdu un trūkstošo atribūtu identificēšanai.
- **Metadata un informācijas arhitektūras pārvaldība:** CoE ir atbildīgs par vienotas klasifikācijas un identifikatoru, nosaukumu un vienību aprakstu standartu sistēmas izveidi, kas ir ļoti svarīga sistēmu integrācijai.



**Attēls 4.3-10 Datu modelēšana un datu kvalitātes pārvaldība ir viens no galvenajiem uzdevumiem CoE**

Datu izcilības centrs (CoE) nav tikai ekspertu grupa, bet gan sistēmisks mehānisms, kas rada jaunu uz datiem balstītu kultūru un nodrošina vienotu pieejumu darbam ar datiem visā uzņēmumā. Kompetenti integrējot modelēšanas procesus kopējā informācijas pārvaldības sistēmā, standartizējot, klasificējot un kontrolējot datu kvalitāti, CoE palīdz uzņēmumiem nepārtrauktī uzlabot savus produktus un uzņēmējdarbības procesus, ātrāk reaģēt uz tirgus izmaiņām un pieņemt pamatotus lēmumus, pamatojoties uz ticamu analīzi.

Šādi centri ir īpaši efektīvi, ja tos apvieno ar mūsdienīgiem DataOps principiem - saskaņā ar gājienu, kas nodrošina nepārtrauktu datu piegādi, automatizāciju un kvalitātes kontroli. Plašāk par DataOps mēs runāsim 8. daļā, nodaļā "Modernās datu tehnoloģijas būvniecības nozarē".

Turpmākajās nodaļās pārīsim no stratēģijas uz praksi - nosacīti "pārveidosim" datu centru: aplūkosim vairākus piemērus, kā notiek uzdevumu parametrizācija, prasību apkopošana un automātiskās validācijas process.



## NODAĻA 4.4.

### PRASĪBU SISTEMATIZĀCIJA UN INFORMĀCIJAS APSTIPRINĀŠANA

#### **Prasību apkopošana un analīze: saziņas pārveidošana strukturētos datos.**

Prasību apkopošana un pārvaldība ir pirmsais solis datu kvalitātes nodrošināšanā. Neraugoties uz digitālo rīku attīstību, lielākā daļa prasību joprojām tiek formulētas nestrukturētā veidā: vēstulēs, sanāksmju protokolos, telefona zvanos un mutiskās diskusijās. Šāds saziņas veids apgrūtina informācijas automatizāciju, apstiprināšanu un atkārtotu izmantošanu. Šajā nodaļā aplūkosim, kā tekstuālās prasības pārvērst formālās struktūrās, nodrošinot, ka biznesa prasības ir pārskatāmas un sistemātiskas.

Gartner pētījumā "Data Quality: Best Practices for Accurate Insights" ("Datu kvalitāte: labākā prakse precīzām atzinīām") uzsvērta datu kvalitātes izšķirošā nozīme veiksmīgu datu un analītikas iniciatīvu īstenošanā [100]. Viņi norāda, ka slikta datu kvalitāte organizācijām izmaksā vidēji vismaz 12,9 miljonus ASV dolāru gadā un ka uzticami, augstas kvalitātes dati ir būtiski, lai izveidotu uz datiem balstītu uzņēmumu.

Strukturētu prasību trūkums novērtētie pie tā, ka viens un tas pats elements (vienība) un tā parametri dažādās sistēmās var tikt saglabāti dažādās variācijās. Tas ne tikai samazina procesu efektivitāti, bet arī rada laika zudumus, informācijas dublēšanos un nepieciešamību atkārtoti apstiprināt datus, pirms tos var izmantot. Rezultātā pat viens izlaidums - pazaudēts parametrs vai viens nepareizi aprakstīts elements - var palēnināt lēmumu pieņemšanu un izraisīt neefektīvu resursu izmantošanu.

*Tā kā nebija naglas, pakava tika pazaudēta.*

*Zirgs tika pazaudēts, jo trūka zemakšas.*

*Zirga trūkuma dēļ tika pazaudēts jātnieks.*

*Tā kā nebija jātnieka, pazuda ziņa.*

*Tā kā trūka vēstījuma, kauja tika zaudēta.*

*Kaujas trūkuma dēļ tika zaudēta kara/valsts.*

*Tas viss tikai tāpēc, ka nebija naglas zemakā.*

- Sakāmvārds [101]

Datu aizpildīšanas un glabāšanas procesa prasību analīze un apkopošana sākas ar visu ieinteresēto personu apzināšanu. Līdzīgi kā vienas naglas zaudēšana izraisa virkni kritisku seku, arī uzņēmējdarbībā vienas ieinteresētās personas zaudējums, vienas neievērotas prasības vai pat viena parametra zudums var būtiski ietekmēt ne tikai atsevišķu uzņēmējdarbības procesu, bet arī visu projekta ekosistēmu un organizāciju kopumā. Tāpēc ir joti svarīgi identificēt pat tos elementus, parametrus un lomas, kas pirmajā brīdī šķiet nenozīmīgi, bet vēlāk var izrādīties kritiski svarīgi uzņēmējdarbības ilgtspējai.

Iedomāsimies, ka uzņēmumam ir projekts, kurā klients izvirza jaunu prasību - "pievienot papildu logu

*"ēkas ziemeļu pusē". Nelielajā procesā "klienta pieprasījums pievienot jaunu logu pašreizējam projektam" piedalās arhitekts, klients, CAD speciālists (BIM), būvdarbu vadītājs, loģistikas vadītājs, ERP -analītikis, kvalitātes kontroles inženieris, drošības inženieris, kontroles vadītājs un īpašuma pārvaldnies.*

*Pat nelielā procesā var būt iesaistīti desmitiem dažādu speciālistu. Katram procesa dalībniekam ir jāizprot prasības tiem speciālistiem, ar kuriem viņi ir saistīti datu līmenī.*

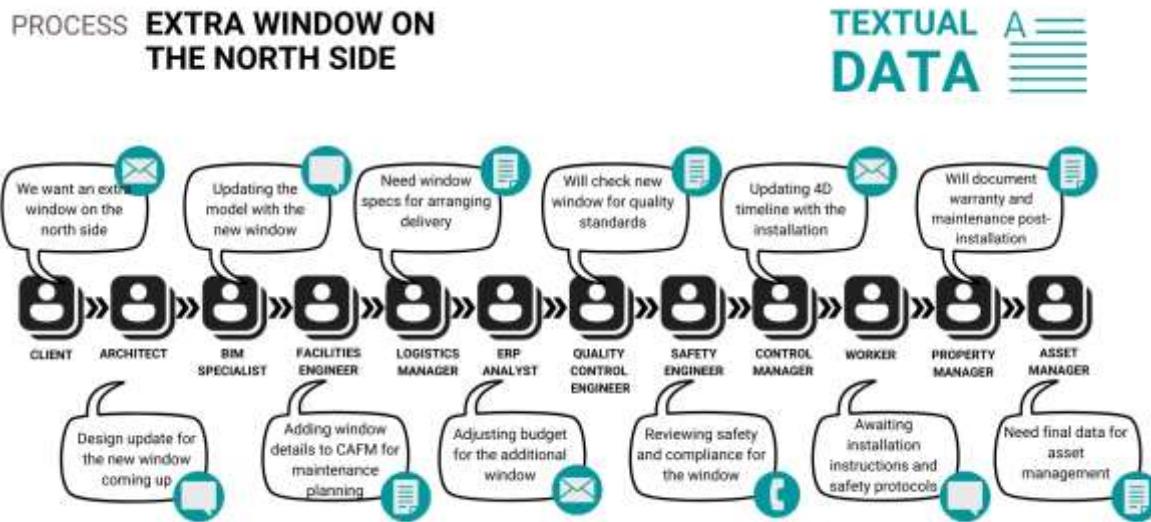
Teksta līmenī ( 4.4-1. attēls) saziņa starp klientu un procesa kēdes speciālistiem notiek šādi:

- ❶ **Klients:** "Esam nolēmuši ziemeļu pusē pievienot papildu logu, lai nodrošinātu labāku apgaismojumu. Vai to var realizēt?"
- ❷ **Arhitekts:** "Protams, es pārstrādāšu projektu, lai iekļautu jauno logu, un nosūtīšu atjauninātus CAD plānus (BIM)".
- ❸ **CAD speciālists (BIM):** "Saņemts jauns projekts. Atjauninu CAD (BIM) modeli ar papildu logu un pēc saskaņošanas ar FEM inženieri nodrošinu precīzu jaunā loga atrašanās vietu un izmērus".
- ❹ **Būvdarbu vadītājs:** "Ir saņemts jauns projekts. Mēs koriģējam 4D uzstādīšanas datumus un informējam visus attiecīgos apakšuzņēmējus".
- ❺ **Iekārtu inženieris (CAFM):** "Es ievadīšu 6D datus par jauno logu CAFM sistēmā, lai nākotnē varētu veikt ēku apsaimniekošanu un tehniskās apkopes plānošanu".
- ❻ **Loģistikas vadītājs:** "Man ir vajadzīgi jaunā loga izmēri un svars, lai organizētu loga piegādi uz objektu."
- ❼ **ERP -analītikis:** "Man ir nepieciešamas 5D budžeta atjaunināšanas tabulas un precīzs loga tips mūsu ERP sistēmā, lai atspoguļotu jaunā loga izmaksas kopējā projekta tāmē."
- ❽ **Kvalitātes kontroles inženieris:** "Kad logu specifikācijas būs gatavas, es pārliecināšos, ka tās atbilst mūsu kvalitātes un materiālu standartiem."
- ❾ **Drošības inženieris:** "Es novērtēšu jaunā loga drošības aspektus, īpašu uzmanību pievēršot atbilstībai un evakuācijai saskaņā ar 8D shēmu".
- ❿ **Kontroles vadītājs:** ", mēs atjaunināsim mūsu 4D laika grafiku, lai atspoguļotu jaunā loga uzstādīšanu, un saglabāsim jaunos datus projekta saturā pārvaldības sistēmā."
- ❻ **Darbinieks (uzstādītājs):** "Nepieciešamas instrukcijas par uzstādīšanu, montāžu un darbu izpildes grafiku. Turklāt, vai ir ieviesti kādi īpaši drošības noteikumi, kas man jāievēro?"
- ❼ **Īpašuma pārvaldnies:** "Pēc uzstādīšanas es dokumentēšu garantijas un tehniskās apkopes informāciju ilgtermiņa pārvaldībai".
- ❽ **aktīvu pārvaldnies:** "Iekārtu inženieri, lūdzu, nosūtiet galīgos datus aktīvu izsekošanai un dzīves cikla pārvaldībai."
- ❾ **Klients:** "Pagaidiet, varbūt es steidzos un logs nebūs vajadzīgs. Varbūt man vajadzētu izveidot balkonu."

Šādos scenārijos, kas notiek bieži, pat nelielas izmaiņas izraisa kēdes reakciju starp vairākām sistēmām un lomām. Šādā gadījumā gandrīz visa saziņa sākotnējā posmā notiek teksta formā: e-pasti, tērzēšanas, sanāksmju protokoli ( 4.4-1. attēls).

Šādā uz tekstu balstītā būvniecības projekta saziņas sistēmā joti svarīga ir visu datu apmaiņas operāciju un pieņemto lēmumu juridiska apstiprināšanas un reģistrēšanas sistēma. Tas ir, lai nodrošinātu, ka katrs lēmums, norādījums vai veiktā izmaiņa ir juridiski derīga un izsekojama, tādējādi

samazinot turpmāku "pārpratumu" risku

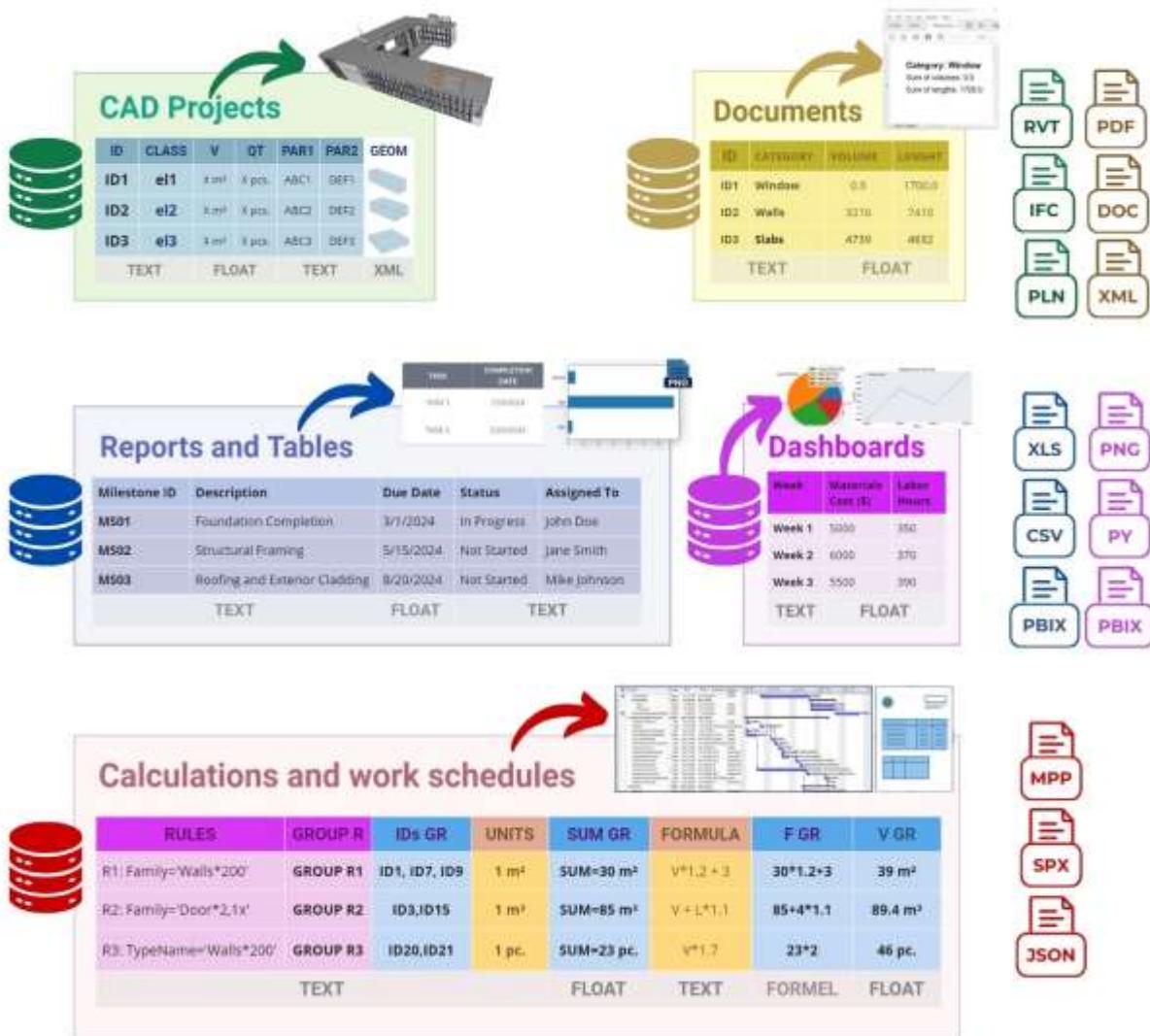


Attēls 4.4-1 Saziņa starp pasūtītāju un izpildītāju projekta sākumposmā bieži vien satur daudzformāta teksta datus.

Juridiskās kontroles un lēmumu apstiprināšanas trūkums attiecīgajās būvniecības projekta sistēmās var radīt nopietnas problēmas visiem iesaistītajiem. Katrs lēmums, rīkojums vai izmaiņa, kas pieņemta bez pienācīgas dokumentācijas un apstiprināšanas, var izraisīt strīdus (un tiesvedību).

Visu lēmumu juridisku konsolidāciju tekstuālā saziņā var nodrošināt tikai ar lielu skaitu parakstītu dokumentu, kas gulstas uz vadības pleciem, kurai ir pienākums reģistrēt visus darījumus. Rezultātā, ja katram dalībniekam par katu darbību ir jāparaksta dokumenti, sistēma zaudē elastību un kļūst par birokrātisku labirintu. Darījumu apstiprinājumu trūkums ne tikai aizkavēs projekta īstenošanu, bet arī var radīt finansiālus zaudējumus un pasliktināt dalībnieku savstarpējās attiecības, līdz pat juridiskām problēmām.

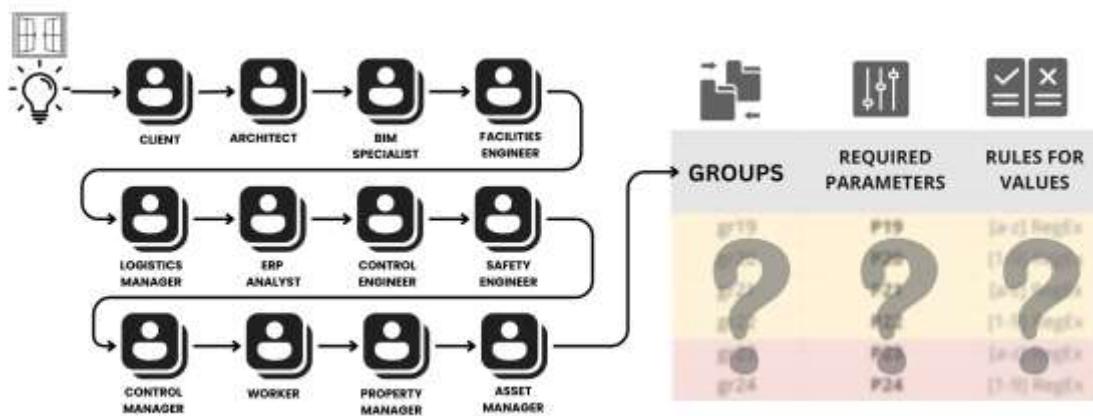
Šāds darījuma apstiprināšanas process, kas parasti sākas ar teksta diskusijām, turpmākajos posmos pakāpeniski pārvēršas par daudzformātu dokumentu apmaiņu (4.4-2. attēls), ievērojami sarežģī saziņu, kas agrāk notika tikai teksta veidā. Bez skaidri definētām prasībām automatizēt šādus procesus, kas ir piepildīti ar daudzformāta datiem un lielu skaitu teksta prasību, kļūst gandrīz neiespējami.



Attēls 4.4-2 Katra sistēma būvniecības uzņēmuma vidē kalpo kā juridiski nozīmīgu dokumentu avots dažados formātos.

Teksta saziņai ir nepieciešams, lai katrs speciālists vai nu iepazītos ar pilnu saraksti, vai regulāri apmeklētu visas sanāksmes, lai saprastu projekta pašreizējo statusu.

Lai pārvarētu šo ierobežojumu, ir nepieciešama pāreja no teksta komunikācijas uz strukturētu prasību modeli. Tas ir iespējams tikai sistemātiski analizējot, vizualizējot procesus un aprakstot mijiedarbību plūsmas diagrammu un datu modeļu veidā ( 4.4-3. attēls). Tāpat kā datu modelēšanā ( 4.3-7. attēls), mēs pārgājām no kontekstuālā-idejas līmeņa uz konceptuālo līmeni, pievienojot dalībnieku izmantotās sistēmas un rīkus un saiknes starp tiem.



Attēls 4.4-3 Lai uzzinātu, kā pārvaldīt un automatizēt validācijas procesu, ir nepieciešams vizualizēt procesus un strukturēt prasības.

Pirmais solis prasību un attiecību sistematizēšanā ir visu saikņu un attiecību vizualizēšana, izmantojot konceptuālās diagrammas. Konceptuālais līmenis ne tikai atvieglos visu procesu dalībnieku izpratni par visu procesu kēdi, bet arī skaidri parādīs, kāpēc un kam dati (un prasības) ir nepieciešami katrā procesa posmā.

### Procesu diagrammas un konceptuālo ietvaru efektivitāte

Lai mazinātu plāsu starp tradicionālo un moderno pieeju datu pārvaldībai, uzņēmumiem ir apzināti jāpāriet no fragmentētiem tekstuāliem aprakstiem uz struktūrētiem procesu attēlojumiem. Datu evolūcijai - no māla plāksnītēm līdz digitālām ekosistēmām - nepieciešami jauni domāšanas rīki. Un viens no šādiem rīkiem ir konceptuālā modelēšana, izmantojot plūsmas diagrammas. Vizuālu diagrammu - plūsmas diagrammu, procesu diagrammu, mijiedarbības diagrammu - izveide ļauj projekta dalībniekiem apzināties, kā viņu darbības un lēmumi ietekmē visu lēmumu pieņemšanas sistēmu.

Ja procesos nepieciešams ne tikai uzglabāt datus, bet arī analizēt vai automatizēt tos, tad jums ir jāsāk risināt konceptuālā un vizuālā prasību slāņa izveides jautājums.

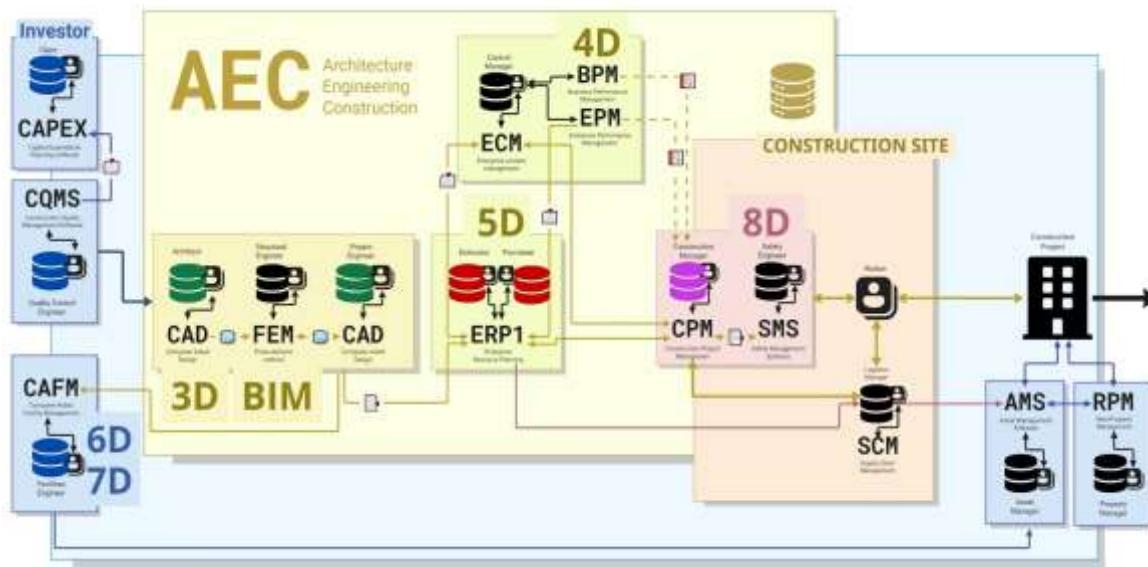
Mūsu piemērā (4.4.4-1. attēls) katrs speciālists var būt daļa no nelielas komandas, bet var būt arī daļa no lielākas nodaļas, kurā ir līdz pat ducim speciālistu, ko kontrolē galvenais vadītājs. Katra nodaļa izmanto specializētu lietojumprogrammu datubāzi (1.2-4. att., piemēram, ERP, CAD, MEP, CDE, ECM, CPM u. c.), kas regulāri tiek papildināta ar ienākošo informāciju, kura nepieciešama dokumentu izveidei, lēmumu juridiskā statusa reģistrēšanai un procesu pārvaldībai.

Darījuma process ir līdzīgs seno vadītāju darbam pirms 4000 gadiem, kad lēmumu juridiskai apstiprināšanai izmantoja māla plāksnītes un papirusu. Mūsdienu sistēmas no to māla un papīra priekštečiem atšķiras ar to, ka modernās metodes papildus ietver teksta informācijas pārvēršanas procesu digitālā formā tālākai automātiskai apstrādei citās sistēmās un rīkos.

Procesa vizualizācijas izveide konceptuālu diagrammu veidā palīdzēs aprakstīt katru soli un mijiedarbību starp dažādām lomām, padarot sarežģīto darba plūsmu skaidru un vienkāršu.

Vizualizējot procesus, tiek nodrošināta procesu loģikas pārredzamība un pieejamība visiem komandas locekļiem.

Tas pats komunikācijas process loga pievienošanai projektam, kas tika aprakstīts teksta, ziņojumu (4.4-1. attēls) un bloku diagrammas veidā, ir līdzīgs konceptuālajam modelim, ko mēs aplūkojām nodaļā par datu modelēšanu (4.4-4. attēls).



**Attēls 4.4-4 Konceptuālā diagrammā projekta dalībnieki ir attēloti kā datubāzes lietotāji, kuru pieprasījumi savieno dažādas sistēmas.**

Lai gan konceptuālās diagrammas ir svarīgs solis, daudzi uzņēmumi aprobežojas tikai ar šo līmeni, uzskatot, ka procesu izpratnei pietiek ar vizuālu diagrammu. Tas rada vadāmības ilūziju: vadītājiem ir vieglāk uz šādas diagrammas uztvert kopainu un saskaņīt saiknes starp dalībniekiem un posmiem. Tomēr šādas shēmas nesniedz skaidru priekšstatu par to, kādi dati ir nepieciešami katram dalībniekam, kādā formātā tie ir jānosūta un kādi parametri un atribūti ir obligāti automatizācijai realizācijai. Konceptuālā plūsmas shēma ir vairāk līdzīga maršruta kartei: tā norāda, kurš ar kuru mijiedarbojas, bet neatklāj, kas šajās mijiedarbībās tiek pārsūtīts.

Pat ja process ir detalizēti aprakstīts konceptuālā līmenī, izmantojot plūsmas diagrammas, tas negarantē tā efektivitāti. Vizualizācija bieži vien atvieglo vadītāju darbu, ļaujot viņiem vieglāk sekot līdzī procesam, izmantojot soli pa solim sagatavotus pārskatus. Tomēr datubāzu inženieriem konceptuālais attēlojums var nebūt pietiekami skaidrs un var nesniegt skaidru izpratni par to, kā procesu īstenot parametru un prasību līmenī.

Pārejot uz sarežģītākām datu ekosistēmām, konceptuālu un vizuālu rīku sākotnējā ieviešana klūst ļoti svarīga, lai nodrošinātu, ka datu procesi ir ne tikai efektīvi, bet arī saskaņoti ar organizācijas stratēģiskajiem mērķiem. Lai pilnībā pārceļtu šo loga pievienošanas procesu ( 4.4-1. attēls) uz datu prasību līmeni, mums ir jāiet vēl dzīlāk un jāpārceļ procesa konceptuālā vizualizācija uz loģisko un fizisko datu, nepieciešamo atribūtu un to robežvērtību līmeni.

## **Strukturētas prasības un RegEx regulārās izteiksmes**

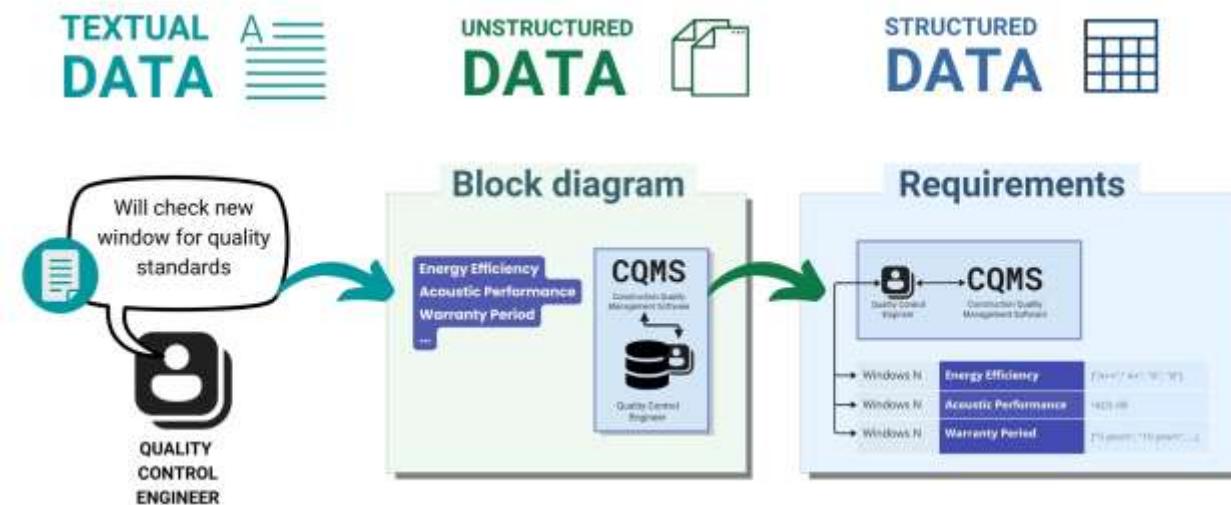
Līdz pat 80% no uzņēmumos radītajiem datiem ir nestukturēti vai daļēji strukturēti [52] - teksts, dokumenti, vēstules, PDF -faili, sarunas. Šādus datus ( 4.4-1. attēls) ir grūti analizēt, pārbaudīt, pārsūtīt starp sistēmām un izmantot automatizācijā.

Lai nodrošinātu pārvaldāmību, pārredzamību un automātisku validāciju, ir nepieciešams teksta un daļēji strukturētas prasības pārvērst precīzi definētos, strukturētos formātos. Strukturēšanas process attiecas ne tikai uz datiem (ko esam detalizēti aplūkojuši šīs grāmatas daļas pirmajās nodalās), bet arī uz pašām prasībām, kuras projekta dalībnieki parasti formulē brīvā teksta formā visā projekta dzīves ciklā, bieži vien neiedomājoties, ka šos procesus var automatizēt.

Tāpat kā mēs jau esam pārveidojuši datus no nestukturētas teksta formas strukturētā formā, prasību darba plūsmā mēs pārveidosim tekstuālās prasības strukturētā "loģiskā un fiziskā slāņa" formātā.

Kā daļa no loga pievienošanas piemēra ( 4.4-1. attēls), nākamas solis ir aprakstīt datu prasības tabulas formā. Mēs strukturēsim informāciju par katru projekta dalībnieku izmantoto sistēmu, norādot galvenos atribūtus un to robežvērtības

Apskatiet, piemēram, vienu no šādām sistēmām ( 4.4-5. attēls) - Būvniecības kvalitātes vadības sistēmu (CQMS), ko izmanto kvalitātes kontroles inženieris no klienta puses. Ar tās palīdzību viņš pārbauda, vai jauns projekta elements - šajā gadījumā "jauns logs" - atbilst noteiktajiem standartiem un prasībām.



Attēls 4.4-5 Teksta prasību pārvēršana tabulas formātā ar entītijas atribūtu aprakstiem atvieglo izpratni citiem speciālistiem.

Kā piemēru aplūkojiet dažas svarīgas prasības attiecībā uz CQMS -sistēmas "logu sistēmas" tipa vienību atribūtiem (4.4-6. attēls): energoefektivitāte, akustiskās īpašības un garantijas termiņš. Katra kategorija ietver noteiktus standartus un specifikācijas, kas jāņem vērā, projektējot un uzstādot logu sistēmas.



Attēlā 4.4-6 Kvalitātes kontroles inženierim jāpārbauda jaunie logu tipa elementi, lai pārbaudītu energoefektivitāti, skaņas izolāciju un garantijas standartus.

Datu prasībām, ko kvalitātes kontroles inženieris norāda tabulas veidā, ir, piemēram, šādas robežvērtības:

- **Logu energoefektivitātes klase** ir no "A++", kas apzīmē augstāko efektivitāti, līdz "B", kas tiek uzskatīta par minimālo pieļaujamo līmeni, un šīs klases ir norādītas pieļaujamo vērtību sarakstā ["A++", "A+", "A", "A", "A", "A", "B"].

- **Logu akustisko izolāciju**, ko mēra decibelos un kas parāda to spēju samazināt ielas troksni, nosaka ar regulāru formulu  $\text{dB}$ .
- Vienības "Loga tips" **atribūta "Garantijas periods"** sākums ir pieci gadi, nosakot šo periodu kā minimālo pieļaujamo, izvēloties produktu; tiek norādītas arī tādas garantijas perioda vērtības kā ["5 gadi", "10 gadi" u. c.] vai loģiskais nosacījums ">5 (gadi)".

Saskaņā ar apkopotajām prasībām noteikto atribūtu ietvaros jaunas logu kategorijas vai klases elementi, kuru novērtējums ir zemāks par "B", piemēram, "C" vai "D", neiztur energoefektivitātes testu. Logu akustiskā izolācija datos vai dokumentos, kas iesniegti kvalitātes nodrošināšanas inženierim, jānorāda ar divciparu skaitli, kam seko postfiks "dB", piemēram, "35 dB" vai "40 dB", un vērtības ārpus šī formāta, piemēram, "9 D B" vai "100 decibeliem", netiks pieņemtas (jo tās neizturēs RegEx virknes paraugu). Garantijas termiņam jāsākas ar vismaz "5 gadi", un logi ar ūsāku garantijas termiņu, piemēram, "3 gadi" vai "4 gadi", neatbilst kvalitātes inženiera aprakstītajām prasībām, kas norādītas tabulas formātā.

Lai validācijas procesā pārbaudītu šādas atribūtu parametru vērtības pret prasību robežvērtībām, mēs izmantojam vai nu atļauto vērtību sarakstu ([A", "B", "C"]], vai vārdnīcas ([A": "H1", "H2"; "B": "W1", W2"]]), loģiskās operācijas (piemēram, ">", "<", "<=", ">=" "==") skaitliskām vērtībām) un regulārās izteiksmes (virknes un teksta vērtībām, piemēram, atribūtā "Acoustic Performance"). Regulārās izteiksmes ir ļoti svarīgs rīks, strādājot ar virknes vērtībām.

**Regulārās izteiksmes** (RegEx) tiek izmantotas programmēšanas valodās, tostarp Python (Re bibliotēka), lai atrastu un modificētu virknes. Regex ir kā detektīvs virķu pasaule, kas spēj precīzi identificēt teksta rakstus tekstā.

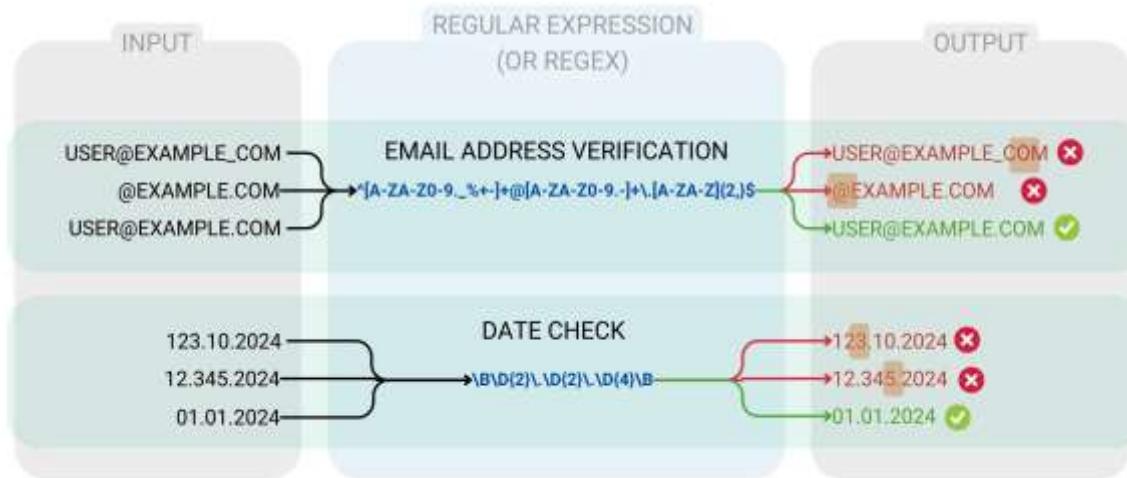
Regulārajās izteiksmēs burti tiek aprakstīti tieši, izmantojot atbilstošos alfabēta simbolus, bet skaitļus var attēlot, izmantojot speciālo rakstzīmi \d, kas atbilst jebkuram ciparam no 0 līdz 9. Lai apzīmētu burtu vai ciparu diapazonu, izmanto kvadrātiekavas, piemēram, [a-z], kas apzīmē jebkuru latīnu alfabēta mazo burtu, vai [0-9], kas atbilst \d. Neskaitļu un neliterāro rakstzīmu apzīmēšanai izmanto attiecīgi \D un \W.

Populāri RegEx lietošanas gadījumi ( 4.4.4-7. attēls):

- **e-pasta adreses pārbaude:** lai pārbaudītu, vai virkne ir derīga e-pasta adrese, varat izmantot "<sup>^</sup> [a-zA-Z0-9.\_%+-]+@[a-zA-Z0-9.-]+\.[a-zA-Z]{2,}\$\$" veidni.
- **Datuma iegūšana:** "\b\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{2}\d{4}\b" veidni var izmantot, lai iegūtu datumu no teksta DD.MM.GGGGGG formātā.
- **tāluņa numuru verifikācija:** lai verificētu tāluņa numurus formātā +49(000)000-0000, modelis izskatās šādi: "\+\d{2}\(\d{3}\)\d{3}-\d{4}".

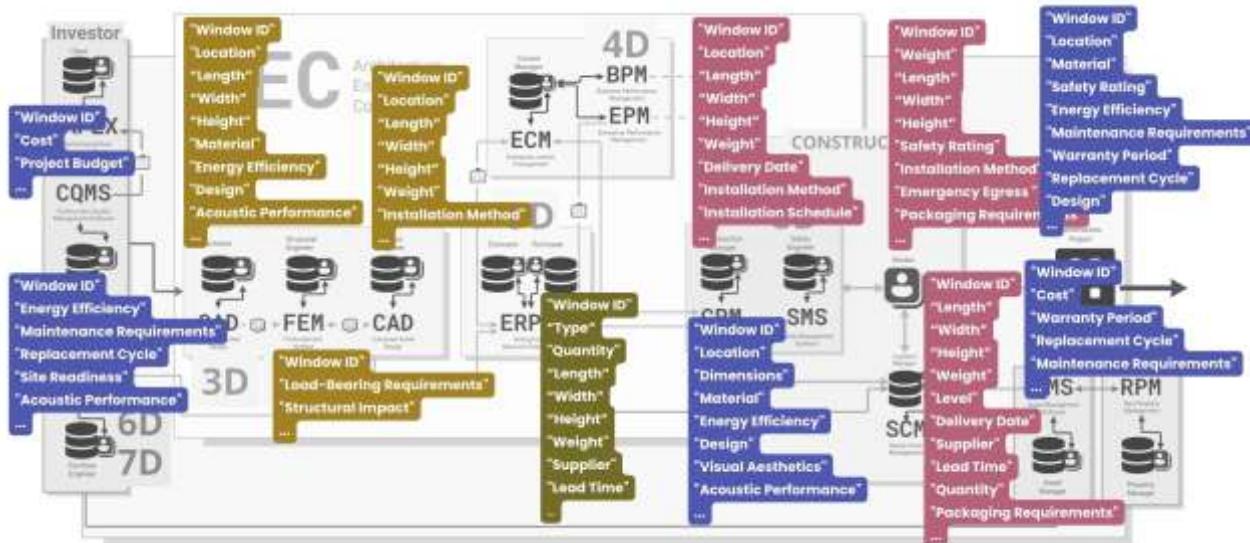
Tulkojot kvalitātes nodrošināšanas inženiera prasības atribūtu un to robežvērtību formātā ( 4.4-6. attēls), mēs esam pārveidojuši tās no sākotnējā teksta formāta (sarunas, vēstules un normatīvie dokumenti) organizētā un strukturētā tabulā, tādējādi ļaujot automātiski pārbaudīt un analizēt jebkurus ienākošos datus (piemēram, jaunus kategorijas Window elementus). Prasību esamība ļauj automātiski noraidīt datus, kas nav pārbaudīti, un pārbaudītus datus automātiski pārsūtīt uz

sistēmām tālākai apstrādei.



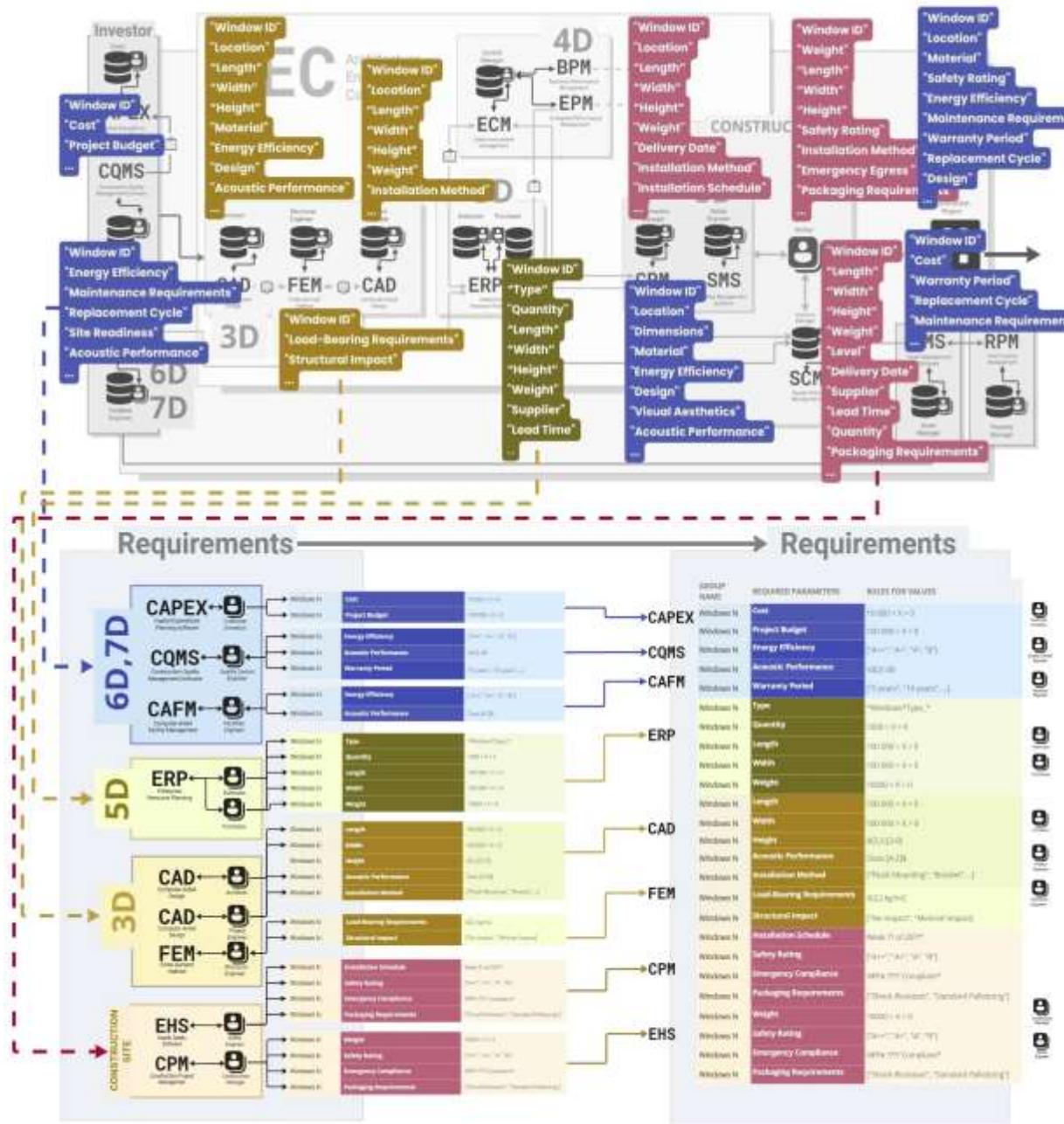
Attēls 4.4.4-7 Regulāro izteiksmju izmantošana ir ārkārtīgi svarīgs rīks teksta datu validēšanas procesā.

Tagad, pārejot no konceptuālā uz loģisko darba ar prasībām līmeni, mēs visu mūsu prasību visu speciālistu prasības mūsu jaunā loga uzstādīšanas procesā ( 4.4-4. attēls ) pārveidosim sakārtotā sarakstā atribūtu formātā un pievienosim šos sarakstus ar nepieciešamajiem atribūtiem mūsu plūsmas diagrammai katram speciālistam ( 4.4-8. attēls ).



Attēls 4.4.4-8 Loģisko procesu līmenī atribūti, ar kuriem strādā katrs speciālists, tiek pievienoti attiecīgajām sistēmām.

Pievienojot visus atribūtus vienai kopīgai procesa tabulai, mēs pārveidojam informāciju, kas iepriekš tika attēlota kā teksts un dialogs konceptuālā līmenī ( 4.4-1. attēls ), strukturētā un sistemātiskā fiziskā līmeņa tabulu formā ( 4.4-9. attēls ).



Attēls 4.4-9 Nestrukturēta speciālistu dialoga pārvēršana strukturētās tabulās palīdz izprast prasības fiziskajā līmenī.

Tagad datu prasības ir jāpaziņo speciālistiem, kas veido informāciju konkrētām sistēmām. Piemēram, ja jūs strādājat CAD datubāzē, pirms sākat elementu modelēšanu, jums jāapkopo visi nepieciešamie parametri, pamatojoties uz datu galīgās izmantošanas scenārijiem. Tas parasti sākas ar ekspluatācijas posmu, kam seko būvlaukums, loģistikas nodaļa, tāmēšanas nodaļa, konstrukciju aprēķinu nodaļa utt. Tikai pēc tam, kad esat ķēmuši vērā visu šo jomu prasības, varat sākt veidot datus - pamatojoties uz apkopotajiem parametriem. Tas jaus jums nākotnē automatizēt datu pārbaudi un pārsūtīšanu visā līmenī.

Ja jaunie dati atbilst prasībām, tie tiek automātiski integrēti uzņēmuma datu ekosistēmā, nonākot tieši pie lietotājiem un sistēmām, kam tie bija paredzēti. Datu pārbaude pēc atribūtiem un to vērtībām nodrošina, ka informācija atbilst nepieciešamajiem kvalitātes standartiem un ir gatava izmantošanai uzņēmuma scenārijos.

Datu prasības ir definētas, un tagad, pirms sākt pārbaudi, ir jāizveido, jāiegūst vai jāsavāc pārbaudāmie dati vai arī jāreģistrē pašreizējais informācijas stāvoklis datubāzēs, lai tos varētu izmantot pārbaudes procesā.

## Datu vākšana pārbaudes procesam

Pirms validācijas uzsākšanas ir svarīgi nodrošināt, lai dati būtu pieejami validācijas procesam piemērotā formā. Tas nenozīmē tikai informācijas pieejamību, bet gan tās sagatavošanu: dati ir jāapkopo un jāpārveido no nestukturētiem, brīvi strukturētiem, teksta un ģeometriskiem formātiem strukturētā formā. Šis process ir detalizēti aprakstīts iepriekšējās nodalās, kur tika aplūkotas dažādu datu veidu pārveidošanas metodes. Visu transformāciju rezultātā ienākošie dati iegūst atvērtu strukturētu tabulu formu (4.1-2. attēls, 4.1-9. attēls, 4.1-13. attēls).

Izmantojot prasības un strukturētas tabulas ar nepieciešamajiem parametriem un robežvērtībām (4.4-9. attēls), mēs varam sākt datu validēšanu - vai nu kā vienotu automatizētu procesu (Pipeline), vai kā pakāpenisku katru ienākošā dokumenta validēšanu.

Lai sāktu pārbaudi, ir nepieciešams vai nu saņemt jaunu failu kā ievadi, vai noteikt datu pašreizējo stāvokli - izveidot momentuzņēmumu vai eksportēt pašreizējos un ienākošos datus, vai izveidot savienojumu ar ārējo vai iekšējo datubāzi. Aplūkojamajā piemērā šāds momentuzņēmums tiek izveidots, automātiski konvertējot CAD datus no strukturētā formātā, kas reģistrēti, piemēram, piektdien, 2024. gada 29. martā, 23:00:00, pēc tam, kad visi projektētāji ir devušies mājās.



Attēls 4.4-10 CAD datubāzes momentuzņēmums (BIM), kurā redzama pašreizējā projekta modeļa pašreizējā versijā esošā atribūtu informācija jaunai "Window" klases vienībai.

Pateicoties reversās inženierijas rīkiem, kas aplūkoti nodalā "CAD datu (BIM) tulkošana strukturētā formā", šo informāciju no dažādiem CAD (BIM) rīkiem un redaktoriem var sakārtot atsevišķās tabulās (4.4-11. attēls) vai apvienot vienā kopīgā tabulā, kas apvieno dažādas projekta daļas (9.1-10. attēls).

Šādā tabulā - datu bāzē - ir redzami logu un durvju unikālie identifikatori (ID atribūts), tipu nosaukumi (TypeName), izmēri (Platums, Garums), materiāli (Material), kā arī energoefektivitātes un akustiskie rādītāji un citas īpašības. Šādu CAD programmā (BIM) aizpildītu tabulu projektēšanas inženieris apkopo no dažādām nodalām un dokumentiem, veidojot projekta informācijas modeli.

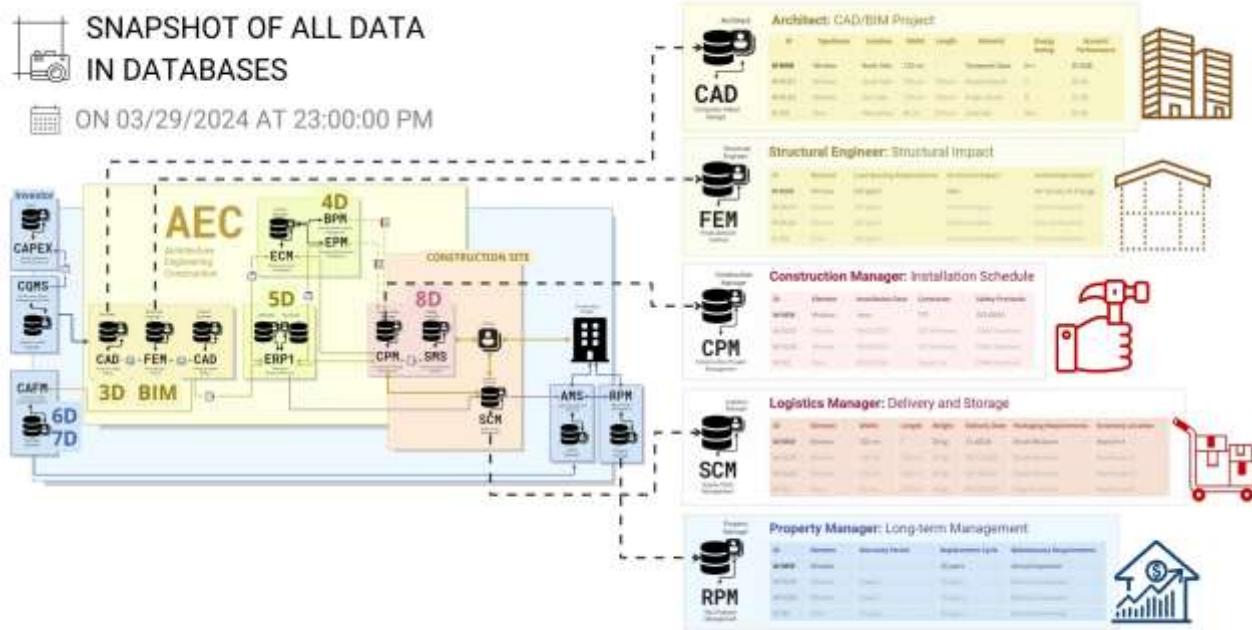


**Attēls 4.4-11 Strukturētie dati no CAD sistēmām var būt divdimensiju tabula ar kolonnām, kas apzīmē elementu atribūtus.**

Reālos CAD (BIM) projektos ir desmitiem vai simtiem tūkstošu elementu (9.1-10. attēls). Elementi CAD formātos tiek automātiski iedaļīti kategorijās pēc tipa un kategorijas - no logiem un durvīm līdz plātnēm, pārsegumiem un sienām. Unikāli identifikatori (piemēram, dzimtā ID, ko automātiski nosaka CAD risinājums) vai tipa atribūti (Type Name, Type, Family) ļauj vienu un to pašu objektu izsekot dažādās sistēmās. Piemēram, jaunu logu ēkas ziemeļu sienā visās attiecīgajās organizācijas sistēmās var unikāli identificēt ar vienotu identifikatoru "W-NEW".

Lai gan vienību nosaukumiem un identifikatoriem jābūt konsekventiem visās sistēmās, ar šīm vienībām saistīto atribūtu un vērtību kopums var ievērojami atšķirties atkarībā no izmantošanas konteksta. Arhitekti, būvinženieri, būvniecības, loģistikas un nekustamā īpašuma uzturēšanas profesionāļi atšķirīgi uzņem vienus un tos pašus elementus. Katrs no viņiem paļaujas uz saviem klasifikatoriem, standartiem un mērķiem: daži skatās uz logu tikai no estētiskā viedokļa, novērtējot tā formu un proporcijas, bet citi uz to raugās no inženiertehniskā vai ekspluatācijas viedokļa, analizējot siltumvadītspēju, uzstādīšanas metodi, svaru vai apkopes prasības. Tāpēc, modelējot datus un aprakstot elementus, ir svarīgi ņemt vērā to izmantošanas daudzveidību un nodrošināt datu konsekvenči, vienlaikus ņemot vērā nozares specifiku.

Katrai uzņēmuma procesu lomai ir specializētas datubāzes ar savu lietotāja saskarni - no projektēšanas un aprēķiniem līdz loģistikai, montāzai un ēkas ekspluatācijai (4.4.4-12. attēls). Katru šādu sistēmu pārvalda profesionāla speciālistu komanda, izmantojot īpašu lietotāja saskarni vai datubāzes pieprasījumus, kur aiz visu lēmumu summas, kas pieņemti par ievadītajām vērtībām ķēdes galā, stāv sistēmas vadītājs vai nodalas vadītājs, kurš atbild par ievadīto datu juridisko derīgumu un kvalitāti saviem partneriem, kas apkalpo citas sistēmas.



Attēls 4.4-12 Vienai un tai pašai vienībai dažādās sistēmās ir viens un tas pats identifikators, bet dažādi atribūti, kas ir svarīgi tikai attiecīgajā sistēmā.

Kad esam organizējuši strukturēto prasību un datu vākšanu logiskā un fiziskā līmenī, mums atliek izveidot procesu, lai automātiski pārbaudītu datus no dažādiem ienākošajiem dokumentiem un dažādām sistēmām, salīdzinot tos ar iepriekš apkopotajām prasībām.

### datu verifikācija un verifikācijas rezultāti

Visi jaunie dati, kas tiek ievadīti sistēmā - dokumenti, tabulas vai datubāzes ieraksti no klienta, arhitekta, inženiera, meistara, loģistikas speciālista vai nekustamā īpašuma pārvaldnika - ir jāapstiprina atbilstoši iepriekš formulētajām prasībām (4.4-9. attēls). Apstiprināšanas process ir ļoti svarīgs: jebkuras klūdas datos var novest pie nepareiziem aprēķiniem, grafika kavējumiem un pat finansiāliem zaudējumiem. Lai līdz minimumam samazinātu šādus riskus, jāorganizē sistemātiska un atkārtojama, iteratīva datu validēšanas procedūra.

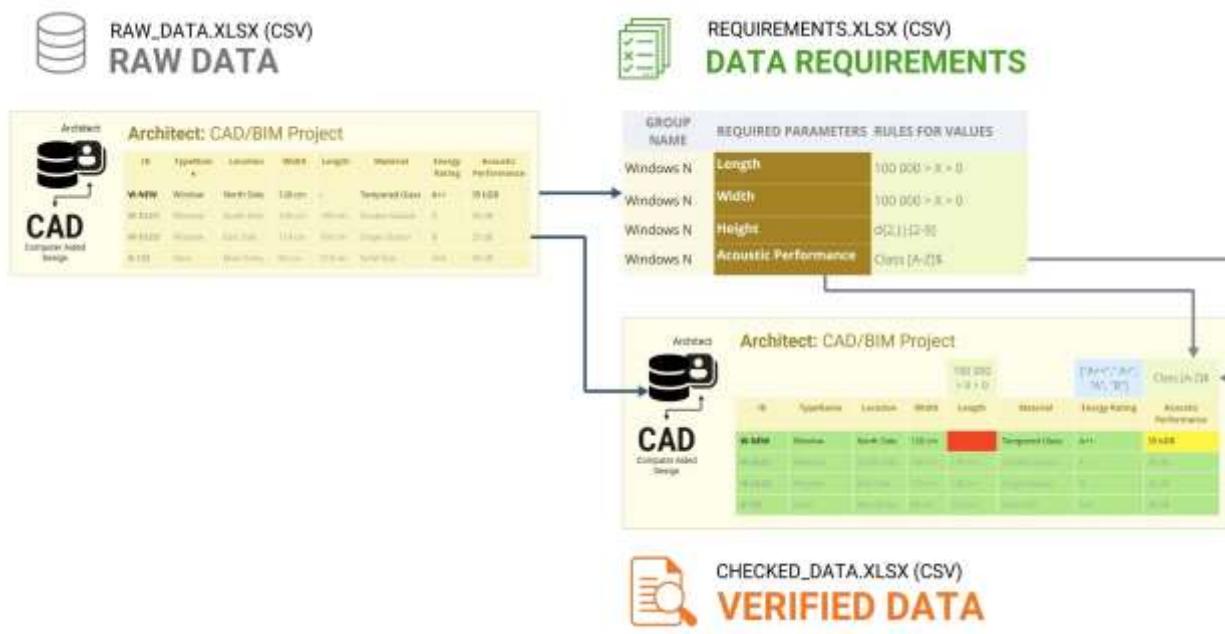
Lai apstiprinātu sistēmā ienākošos jaunos datus - nestukturētus, teksta vai ģeometriskus -, tie ir jākonvertē brīvi strukturētā vai strukturētā formātā. Pēc tam validācijas procesā dati jāpārbauda, salīdzinot tos ar pilnu nepieciešamo atribūtu sarakstu un to atļautajām vērtībām.

Dažādu datu veidu: teksta, attēlu, PDF - dokumentu un jauktu CAD (BIM) datu pārveidošana strukturētā formā tika detalizēti aplūkota nodaļā "Datu pārveidošana strukturētā formā".

Piemērs ir tabula, kas iegūta no CAD (BIM) projekta (4.4.4-11. attēls). Tā ietver daļēji strukturētus

ģeometriskos datus un strukturētu atribūtu informāciju par projekta vienībām ( 3.1-14. att.) - piemēram, elementu no klases "Windows".

Lai veiktu validāciju, mēs saīdzinām atritūtu vērtības ( 4.4.4-11. attēls) ar atsauces robežvērtībām, ko eksperti ir definējuši prasību veidā ( 4.4.4-9. attēls). Gala saīdzinājuma tabulā ( 4.4-13. att.) būs skaidri redzams, kuras vērtības ir pieņemamas un kuras ir jākoriģē, pirms datus var izmantot ārpus CAD lietojumprogrammām (BIM).



Attēls 4.4.4-13 Galīgajā validācijas tabulā ir izceltas tās jaunās klases "Windows" vienības atritūtu vērtības, kurām jāpievērš uzmanība.

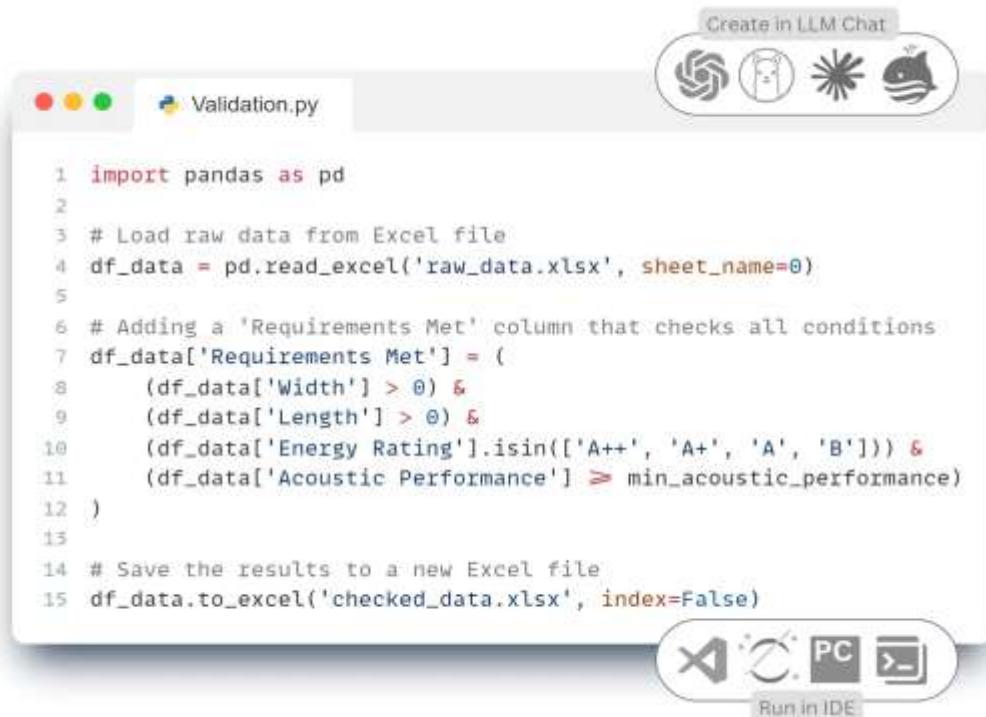
Īstenojot līdzīgu risinājumu, izmantojot Pandas bibliotēku, ko aprakstījām iepriekš nodaļā "Pandas: neaizstājams rīks datu analīzei", mēs apstiprināsim datus no tabulas datnes, kas iegūti no CAD datnes (BIM) (RVT, IFC, DWG, NWS, DGN) (4.4-11. attēls), izmantojot prasības no citas tabulas prasību datnes (4.4-9. attēls).

Lai iegūtu kodu, mums LLM uzvednē ir jāapapraksta, ka mums ir nepieciešams ielādēt datus no faila **raw\_data.xlsx** (pilns datu kopums no CAD datubāzes (BIM)), **pārbaudīt** tos un saglabāt rezultātu jaunā failā **checked\_data.xlsx** ( 4.4-13. attēls).

❷ legūsim kodu, izmantojot LLM, neminot Pandas bibliotēku:

Uzrakstiet kodu, lai pārbaudītu tabulu no raw\_data.xlsx faila un validētu tos, izmantojot šādus validācijas noteikumus: "Platums" un "Garums" slejas vērtības ir lielākas par nulli, "Enerģijas kategorija" ir iekļauta sarakstā ["A++", "A+", "A", "B"] un "Akustiskās īpašības" ir mainīgais lielums, ko norādīsim vēlāk - ar pievienotu pēdējo validācijas kolonnu, un saglabājet galīgo tabulu jaunā Excel failā checked\_data.xlsx.

- LLM atbildē tiks aprakstīts ūss Python koda piemērs, kuru var pilnveidot un papildināt ar turpmākajiem uzaicinājumiem :



```

1 import pandas as pd
2
3 # Load raw data from Excel file
4 df_data = pd.read_excel('raw_data.xlsx', sheet_name=0)
5
6 # Adding a 'Requirements Met' column that checks all conditions
7 df_data['Requirements Met'] = (
8     (df_data['Width'] > 0) &
9     (df_data['Length'] > 0) &
10    (df_data['Energy Rating'].isin(['A++', 'A+', 'A', 'B'])) &
11    (df_data['Acoustic Performance'] >= min_acoustic_performance)
12 )
13
14 # Save the results to a new Excel file
15 df_data.to_excel('checked_data.xlsx', index=False)

```

**Attēls 4.4-14 LLM- modeļa ģenerētais kods pārbauda konvertētā CAD (BIM ) projekta atbilstību atribūtu prasībām robežu vērtību veidā.**

LLM valodas modeļa ģenerēto kodu, var izmantot jebkurā populārā IDE vai tiešsaistes rīkā: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populāros tiešsaistes rīkos Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

Izpildot kodu ( 4.4.4-14. attēls), redzams, ka "vienības elementi" W-OLD1, W-OLD2, D-122 (un citi elementi) no CAD datubāzes (BIM) atbilst atribūtu prasībām: platumis un garums ir lielāks par nulli, un energoefektivitātes klase ir viena no saraksta vērtībām "A++", "A", "B", "C" ( 4.4-15. attēls).

Mums vajadzīgais un nesen pievienotais W-NEW elements, kas atbild par jauno elementu klasī "Window" ziemeļu pusē, neatbilst prasībām (atribūts "Requrments Met"), jo tā garums ir nulle (vērtība "0,0" tiek uzskatīta par nepieņemamu saskaņā ar mūsu noteikumu "Width">>0) un tajā nav norādīta energoefektivitātes klase.


**CHECKED\_DATA.XLSX (CSV)**  
**VERIFIED DATA**

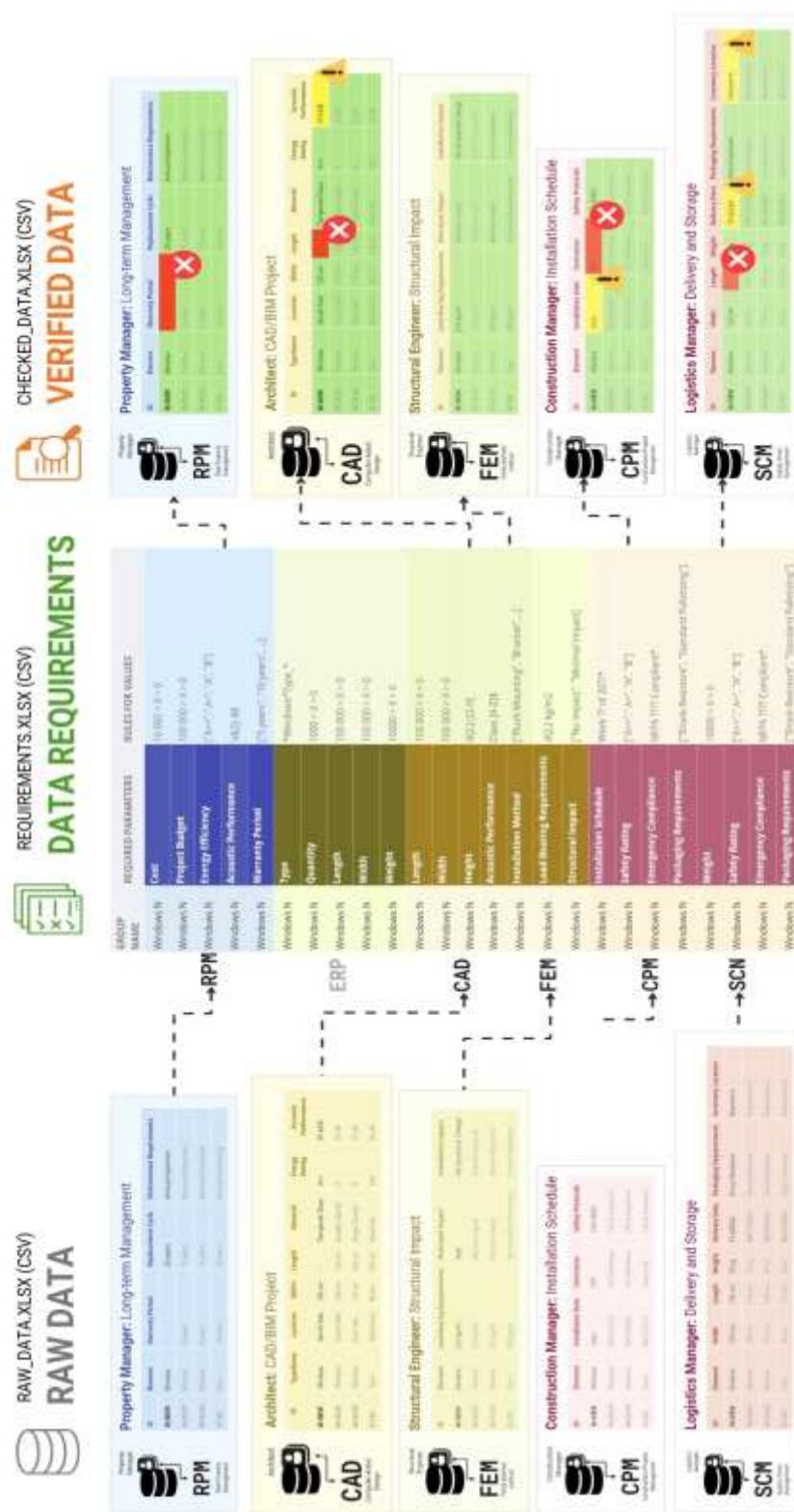
ID	TypeName	Location	Width	Length	Material	Energy Rating	Acoustic Performance	Requirements Met	
0	W-NEW	Window	North Side	120	0.0	Tempered Glass		35	False
1	W-OLD1	Window	South Side	100	140.0	Double Glazed	A++	30	True
2	W-OLD2	Window	East Side	110	160.0	Single Glazed	B	25	True
3	D-122	Door	Main Entry	90	210.0	Solid Oak	B	30	True

**Attēls 4.4-15 Pārbaude identificē vienības, kas nav izgājušas verifikācijas procesu, un rezultātiem pievieno jaunu atribūtu ar vērtībām "False" vai "True".**

Līdzīgi mēs pārbaudām visu projekta elementu (vienību) un nepieciešamo atribūtu atbilstību katrai no sistēmām, tabulām vai datubāzēm visos datos, ko saņemam no dažādiem speciālistiem ( 4.4-1. attēls) loga pievienošanas procesā projektam.

Galīgajā tabulā ir ērti vizualizācijai krāsās izcelt pārbaudes rezultātus: zaļā krāsā tiek atzīmēti atribūti, kas veiksmīgi izturējuši pārbaudi, dzeltenā krāsā - vērtības ar nekritiskām novirzēm, bet sarkanā krāsā - kritiskās neatbilstības ( 4.4-16. att.).

Apstiprināšanas rezultātā ( 4.4.4-16. attēls) tiek iegūts uzticamo un apstiprināto elementu saraksts ar to identifikatoriem, kuru atbilstība atribūtu prasībām ir pārbaudīta. Apstiprinātie elementi sniedz pārliecību, ka šie elementi atbilst noteiktajiem standartiem un specifikācijām visām sistēmām, kas iesaistītas Window klases vai jebkuras citas klases elementu pievienošanas procesā (vairāk par datu validēšanas automatizēšanu un automatizēta ETL procesa izveidi mēs aplūkosim nodaļā "ETL un datu validēšanas automatizēšana").



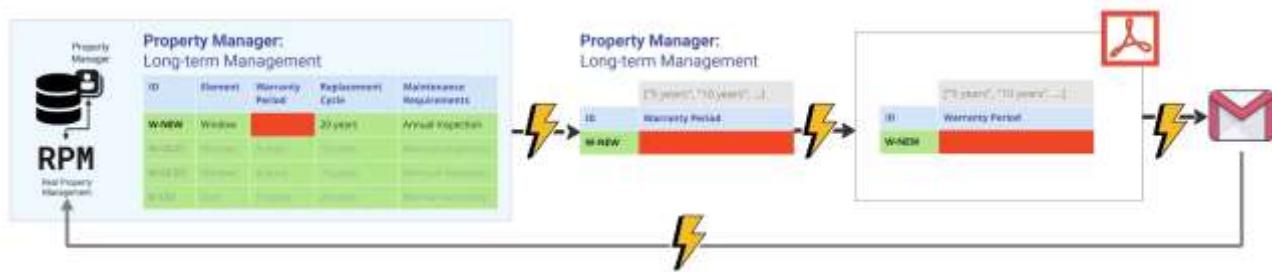
**Attēls 4.4-16 Visu sistēmu pārbaudes rezultāts** Iauj noteikt, kuri dati neatbilst uzņēmuma prasībām.

Vienībām, kas ir veiksmīgi validētas, parasti nav jāpievērš liela uzmanība. Tās bez šķēršļiem pāriet uz nākamajiem apstrādes un integrācijas posmiem citās sistēmās. Atšķirībā no "kvalitatīviem" vienumiem vislielāko interesiju izraisa tie vienumi, kas nav izturējuši validāciju. Informācija par šādām novirzēm ir ļoti svarīga: tā ir jāpaziņo ne tikai tabulas pārskatos, bet arī izmantojot dažādus vizualizācijas rīkus. Verifikācijas rezultātu grafisks attēlojums palīdz ātri novērtēt datu kvalitātes vispārējo stāvokli, identificēt problemātiskās jomas un nekavējoties veikt koriģējošus vai koriģējošus pasākumus.

## Verifikācijas rezultātu vizualizācija

Vizualizācija ir būtisks rīks pārbaudes rezultātu interpretēšanai. Papildus ierastajām kopsavilkuma tabulām tajā var iekļaut paneļus, diagrammas un automātiski ģenerētus PDF dokumentus, kas grupē projekta elementus pēc to pārbaudes statusa. Krāsu kodiem var būt atbalstoša nozīme: zaļa krāsa var norādīt uz elementiem, kas ir veiksmīgi pārbaudīti, dzeltena - uz elementiem, kam jāpievērš papildu uzmanība, bet sarkana - uz elementiem, kuros ir kritiskas klūdas vai trūkst svarīgu datu.

Mūsu piemērā (4.4-1. attēls) mēs analizējam datus no katras sistēmas soli pa solim: no CAD (BIM) un īpašuma pārvaldības līdz logistikai un uzstādīšanas grafikiem (4.4-16. attēls). Pēc audita katram speciālistam automātiski tiek ģenerēti individuāli brīdinājumi vai atskaites dokumenti, piemēram, PDF formātā (4.4-17. att.). Ja dati ir pareizi, speciālists saņem ūsu ziņojumu: "Paldies par sadarbību". Ja tiek konstatētas neatbilstības, tiek nosūtīts detalizēts ziņojums ar šādu formulējumu: "Šajā dokumentā uzskaitīti elementi, to identifikatori, atribūti un vērtības, kuru atbilstība nav pārbaudīta".



Attēls 4.4-17 Validēšana un automātiskā pārskatu ģenerēšana paātrina datu trūkumu atklāšanas un izpratnes procesu speciālistam, kas veido datus.

Pateicoties automatizētajam validēšanas procesam - tiklīdz tiek atklāta klūda vai datu nepilnība, personai, kas ir atbildīga par attiecīgo vienību un to atribūtu izveidi vai apstrādi, tiek nosūtīts tūlītējs paziņojums tērzēšanas ziņojuma, e-pasta vai PDF -dokumenta veidā (4.4-18. attēls), norādot to elementu un atribūtu aprakstus, kuri nav validēti.



Attēls 4.4-18 Automātiskie pārbaudes ziņojumi atvieglo kļūdu izpratni un paātrina darbu, aizpildot projekta datus.

Piemēram, ja īpašuma pārvaldības sistēmā (pēc strukturēšanas) nonāk dokuments ar nepareizi aizpildītu atribūtu "Garantijas termiņš", īpašuma pārvaldnieks saņem brīdinājumu ar sarakstu, kurā norādīti pārbaudāmie un labojamie atribūti.

Līdzīgi, ja instalēšanas grafikā vai loģistikas datos būs nepilnības, tiks automātiski sagatavots ziņojums un, piemēram, attiecīgajam speciālistam tiks nosūtīts čata paziņojums vai e-pasta vēstule ar pārbaudes rezultātiem.

Papildus PDF -dokumentiem un grafikiem ar rezultātiem ir iespējams izveidot panelus un interaktīvus 3D -modeļus ( 7.1-6. attēls, 7.2-12. attēls), kuros izcelti elementi ar trūkstošiem atribūtiem, ļaujot lietotājiem vizuāli izmantot elementu 3D ģeometrijas, lai filtrētu un novērtētu šo elementu kvalitāti un pilnību projektā.

Vizualizējot pārbaužu rezultātus automātiski generētu dokumentu, grafiku vai paneļu veidā, tiek ievērojami vienkāršota datu interpretācija un veicināta efektīva saziņa starp projekta dalībniekiem.

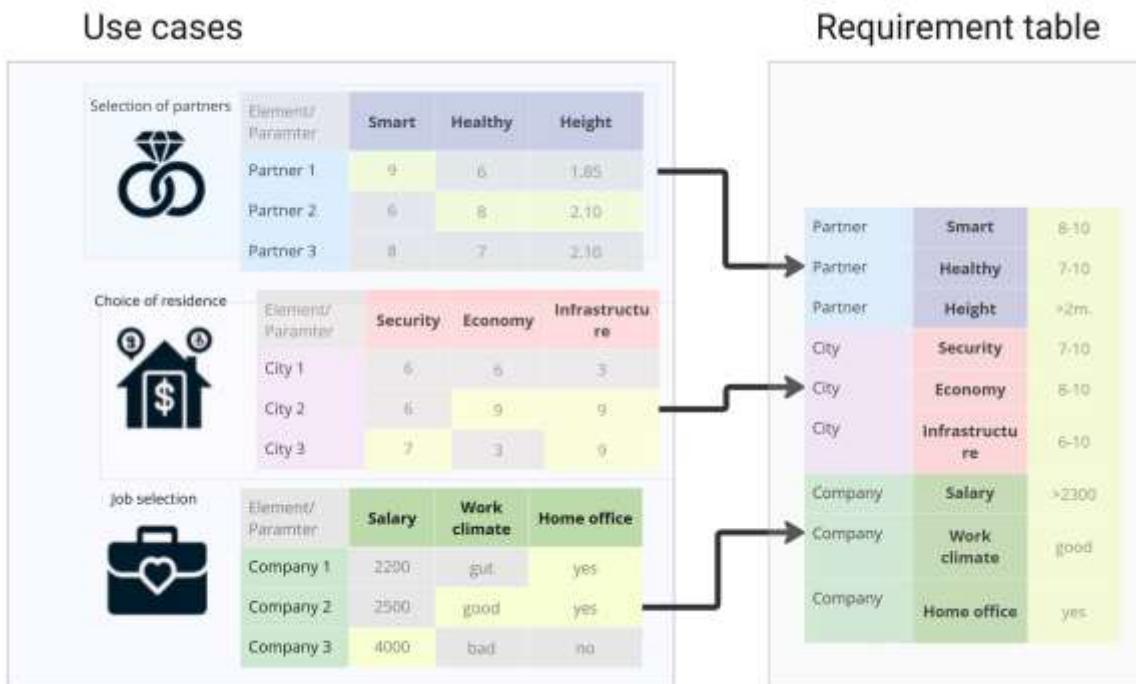
Automātiskas datu pārbaudes procesu no dažādām sistēmām un informācijas avotiem var salīdzināt ar informētu lēmumu pieņemšanu ikdienas dzīvē. Tāpat kā būvniecības nozares uzņēmumi ļem vērā daudzus mainīgos lielumus - no ievades datu ticamības līdz to ietekmei uz projekta īstenošanas laiku, izmaksām un kvalitāti -, arī cilvēks, pieņemot svarīgus lēmumus, piemēram, izvēloties dzīvesvietu, izvērtē virknī faktoru: transporta pieejamību, infrastruktūru, izmaksas, drošību, dzīves kvalitāti. Visi šie apsvērumi veido kritēriju sistēmu, kas ir pamatā galīgajiem lēmumiem, kuri veido mūsu dzīvi.

## Datu kvalitātes pārbaužu salīdzinājums ar cilvēka dzīves vajadzībām

Neraugoties uz datu kvalitātes kontroles metožu un rīku pastāvīgo attīstību, informācijas atbilstības pamatprincips paliek nemainīgs. Šis princips ir iestrādāts nobriedušas pārvaldības sistēmas pamatā gan uzņēmējdarbībā, gan ikdienas dzīvē.

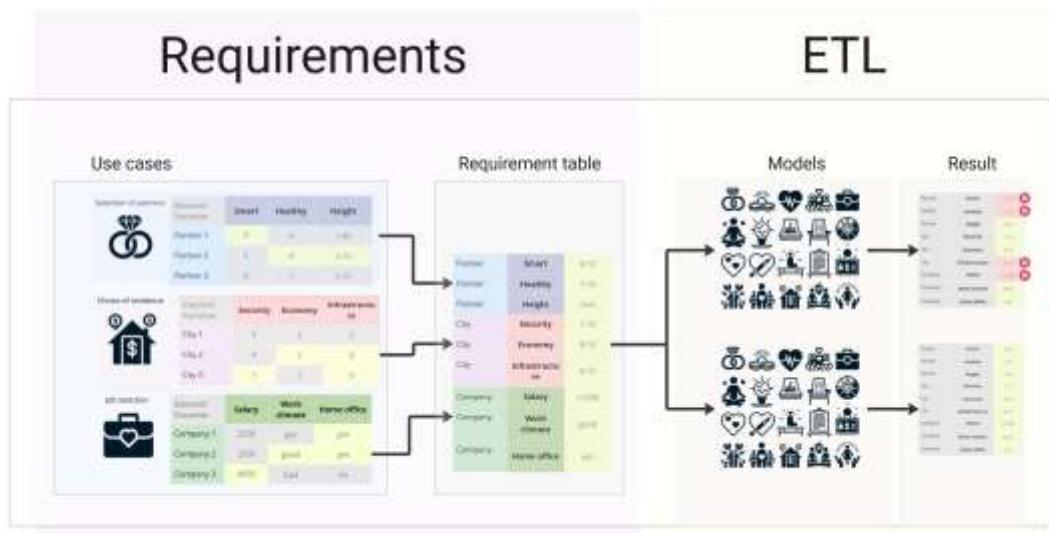
Iteratīvais datu validēšanas process ir līdzīgs lēmumu pieņemšanas procesam, ar ko ikviens saskaras ikdienā. Abos gadījumos mēs paļaujamies uz pieredzi, datiem un jaunu informāciju, kad tā kļūst pieejama. Un arvien vairāk dzīves un profesionālo lēmumu - no stratēģiskiem līdz ikdienas lēmumiem - tiek pieņemti, pamatojoties uz datiem.

Piemēram, izvēloties dzīvesvietu vai dzīves partneri, mēs intuitīvi domās veidojam kritēriju un īpašību tabulu, pēc kuras salīdzinām alternatīvas ( 4.4-19. attēls). Šie raksturlielumi - vai tās būtu cilvēka personiskās īpašības vai īpašuma parametri - ir atribūti, kas ietekmē galīgo lēmumu.



Attēls 4.4-19 Dzīvesvietas, darba vai partnerattiecību izvēle ir balstīta uz individuālām atribūtu prasībām.

Strukturētu datu izmantošana un formalizēta pieeja prasību aprakstīšanai ( 4.4.4-20. attēls) palīdz izdarīt pamatošāku un informētāku izvēli gan profesionālajā, gan personīgajā dzīvē.

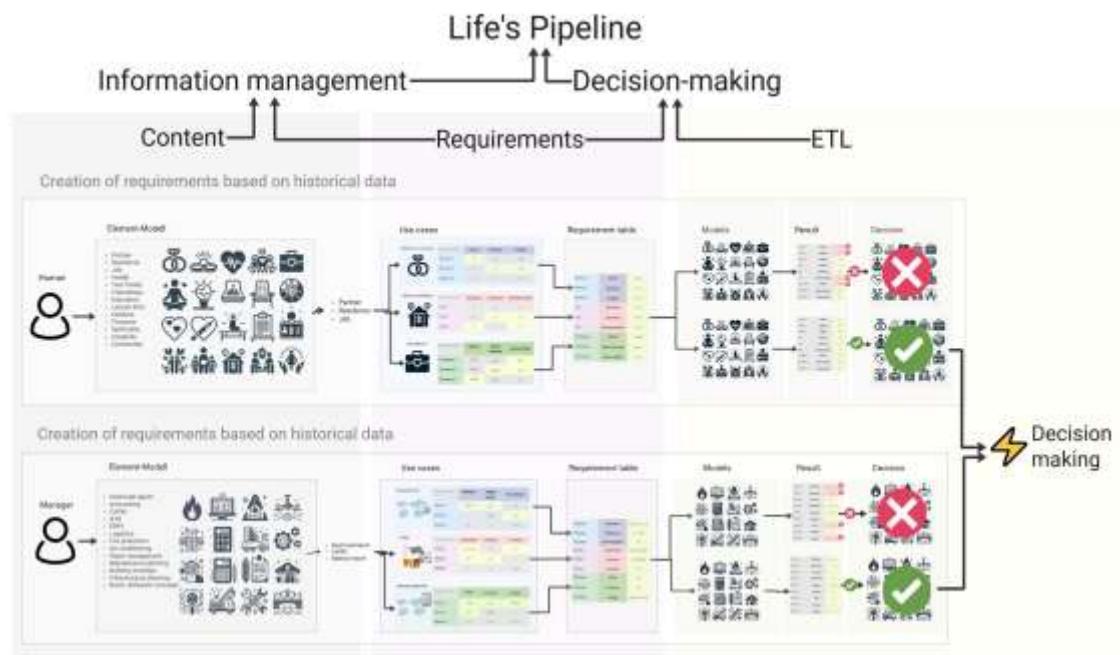


Attēls 4.4-20 Prasību formalizēšana palīdz sistematizēt dzīves un biznesa lēmumu uztveri.

Uz datiem balstīta lēmumu pieņemšanas pieeja nav tikai biznesa rīks. Tā ir nevainojami integrēta arī ikdienas dzīvē, ievērojot kopējos datu apstrādes soļus ( 4.4-21. attēls), kas līdzīgi ETL procesam (Extract, Transform, Load), par kuru mēs jau runājām šīs daļas sākumā, strukturējot datus, un kuru mēs detalizēti aplūkosim uzdevumu automatizācijas kontekstā septītajā grāmatas daļā:

- **Dati kā pamats (izvilkums):** Jebkurā jomā - gan darbā, gan personīgajā dzīvē - mēs apkopojam informāciju. Uzņēmējdarbībā tie var būt pārskati, skaitļi, tirgus dati; personīgajā dzīvē tā var būt personīgā pieredze, tuvinieku padomi, atsauksmes, novērojumi.
- **Vērtēšanas kritēriji (Transform):** savākto informāciju interpretē, pamatojoties uz iepriekš noteiktiem kritērijiem. Darbā tie ir darbības rādītāji (KPI), budžeta ierobežojumi un normas; personīgajā dzīvē - tādi parametri kā cena, ērtība, uzticamība, harizma utt.
- **Prognozes un riska analīze (Load):** pēdējais posms ietver lēmuma pieņemšanu, pamatojoties uz pārveidoto datu analīzi un iespējamo seku salīdzināšanu. Tas ir līdzīgi kā uzņēmējdarbības procesos, kur dati iziet cauri uzņēmējdarbības loģikai un riska filtram.

Lēmumi, ko mēs pieņemam - sākot ar triviālām izvēlēm, piemēram, ko ēst brokastīs, un beidzot ar svarīgiem dzīves notikumiem, piemēram, karjeras vai dzīves partnera izvēli, - pēc būtības ir datu apstrādes un novērtēšanas rezultāts.



**Attēls 4.4-21 Uzņēmējdarbība un dzīve kopumā ir virkne uz datiem balstītu lēmumu, kur galvenais faktors ir lēmumu pieņemšanā izmantoto datu kvalitāte.**

Viss mūsu dzīvē ir savstarpēji saistīts, un tāpat kā dzīvie organismi, tostarp cilvēki, ievēro dabas likumus, attīstās un pielāgojas mainīgajiem apstākļiem, arī cilvēka procesi, tostarp tas, kā mēs vācam un analizējam datus, atspoguļo šos dabas principus. Ciešā saikne starp dabu un cilvēka darbību apliecinā ne tikai mūsu atkarību no dabas, bet arī mūsu vēlmi piemērot miljoniem gadu evolūcijas gaitā pilnveidotos likumus, lai radītu datu arhitektūras, procesus un sistēmas lēmumu pieņemšanai.

Jaunās tehnoloģijas, īpaši būvniecībā, ir spilgts piemērs tam, kā cilvēce, iedvesmojoties no dabas, ik pa laikam rada labākus, ilgtspējīgākus un efektīvākus risinājumus.

## Nākamie soļi: datu pārvēršana precīzos aprēķinos un plānos

Šajā daļā mēs aplūkojām, kā nestukturētus datus pārvērst strukturētā formātā, izstrādāt datu modeļus un organizēt procesus informācijas kvalitātes pārbaudei būvniecības projektos. Datu pārvaldība, standartizācija un klasifikācija ir fundamentāls process, kam nepieciešama sistemātiska pieeja un skaidra izpratne par uzņēmējdarbības prasībām. Šajā daļā aplūkotie paņēmieni un rīki ļauj veikt uzticamu integrāciju starp dažādām sistēmām visā objekta dzīves ciklā.

Apkopojoš šo daļu, izcelsim galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās pieejas ikdienas uzdevumos:

- Sāciet ar prasību sistematizēšanu
- Izveidojiet atribūtu un parametru reģistru galvenajiem projektu un procesu elementiem.
- Katra atribūta robežvērtību dokumentēšana

- Vizualizēt procesus un attiecības starp klasēm, sistēmām un atribūtiem, izmantojot diagrammas (piemēram, Miro, Canva, Visio).

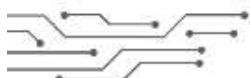
#### **■ Automatizēt datu konvertēšanu**

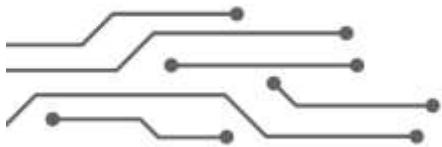
- Pārbaudiet, kurus no procesos bieži izmantotajiem dokumentiem var digitalizēt, izmantojot OCR bibliotēkas, un pārvērst tabulārajā formā.
- Pārbaudiet reversās inženierijas rīkus, lai iegūtu datus no CAD (BIM).
- Izmēģiniet iestatīt automātisku datu izgūšanu no dokumentiem vai formātiem, kurus bieži izmantojat darbā, uz tabulas formu.
- Iestatīt automātisku konvertēšanu starp dažādiem datu formātiem

#### **■ Izveidot zināšanu bāzi kategorizēšanai**

- Izstrādāt iekšējo vai izmantot esošo elementu klasifikatoru, kas atbilst nozares standartiem.
- dokumentēt dažādu klasifikācijas sistēmu savstarpējo saistību.
- Apspriediet ar savu komandu tematu par vienotas elementu identifikācijas un viennozīmīgas klasifikācijas sistēmas izmantošanu.
- Sāciet veidot procesu, lai automātiski validētu datus - gan tos, ar kuriem strādājat komandā, gan tos, kuri tiek nosūtīti ārējām sistēmām.

Izmantojot šīs pieejas, varat ievērojami uzlabot datu kvalitāti un vienkāršot to turpmāko apstrādi un pārveidošanu. Turpmākajās šīs grāmatas daļās aplūkosim, kā jau struktūrētus un sagatavotus datus izmantojot automatizētiem aprēķiniem, izmaksu aplēsēm, grafiku sastādīšanai un būvniecības projektu vadībai.





## V DAĻA

# **IZMAKSU UN LAIKA APRĒĶINI: DATU IEKLAUŠANA BŪVΝIECĪBAS PROCESOS.**

Piektā daļa ir veltīta praktiskiem datu izmantošanas aspektiem, lai optimizētu izmaksu aprēķinu un būvniecības projektu plānošanu. Detalizēti tiek analizēta uz resursiem balstīta metode izmaksu aprēķināšanai un aprēķinu procesu automatizācija. Tieki aplūkotas metodes automatizētai daudzuma aprēķinu iegūšanai no CAD (BIM) modeļiem un to integrācija ar aprēķinu sistēmām. Tieki pētītas 4D un 5D modelēšanas tehnoloģijas laika grafiku plānošanai un būvniecības izmaksu pārvaldībai, kā arī konkrēti to piemērošanas piemēri. Tieki analizēti paplašinātie informācijas slāņi 6D -8D, kas nodrošina integrētu pieejumu nekustamā īpašuma objektu ilgtspējības, ekspluatācijas un drošības novērtēšanai. Detalizēti aplūkotas metodes, kā aprēķināt būvniecības projektu oglekļa pēdas nospiedumu un ESG -indikatorus mūsdienu vides prasību un standartu kontekstā. Kritiski tiek izvērtētas tradicionālo ERP un PMIS sistēmu iespējas un ierobežojumi būvniecības procesu pārvaldībā, analizējot to ietekmi uz cenu veidošanas pārredzamību. Tieki prognozētas perspektīvas pārejai no slēgtiem risinājumiem uz atvērtiem standartiem un elastīgiem datu analīzes rīkiem, kas spēj nodrošināt lielāku būvniecības procesu efektivitāti.

## NODĀLA 5.1.

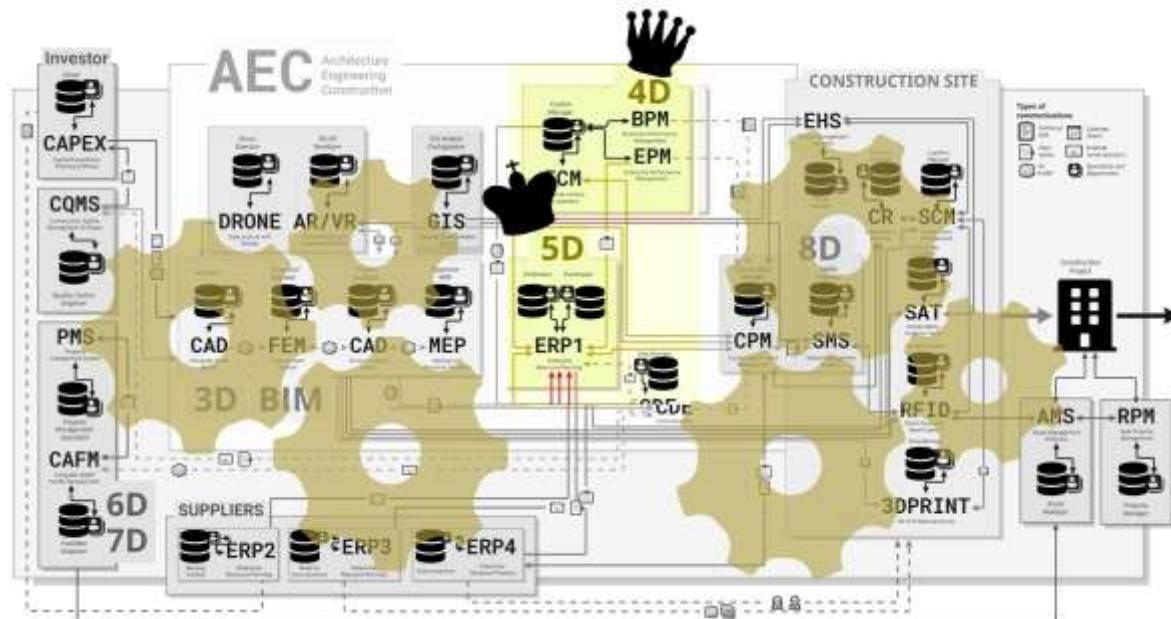
### BŪVΝIECĪBAS PROJEKTU IZMAKSU APRĒĶINI UN TĀMES.

#### **Būvniecības pamati: daudzuma, izmaksu un laika aplēses**

Starp daudzajiem uzņēmējdarbības procesiem, kas nosaka uzņēmuma ilgtspēju būvniecības nozarē, īpaši svarīgi - tāpat kā pirms tūkstošiem gadu - ir procesi, kas ļauj precīzi novērtēt elementu skaitu, projekta izmaksas un piegādes laiku ( 5.1-1. attēls).

Rakstības attīstību noteica vairāki faktori, tostarp nepieciešamība reģistrēt saimnieciskos darījumus, tirdzniecību un resursu pārvaldību agrīnajās sabiedrībās. Pirmie juridiski nozīmīgie dokumenti - māla plāksnītes ar materiālu izmaksu un darbaspēka izmaksu aprēķiniem - tika izmantoti tirdzniecībā un būvniecībā. Šajās plāksnītēs tika fiksētas pušu saistības būvju celtniecībā, un tās tika saglabātas kā liecības par līgumiem un naudas un tirdzniecības attiecībām.

Gadu tūkstošiem ilgi aprēķinu veikšanas pieeja ir saglabājusies gandrīz nemanīga: aprēķini tika veikti manuāli, paļaujoties uz aprēķinu inženiera pieredzi un intuīciju. Tomēr līdz ar modulāro ERP sistēmu un CAD rīku parādīšanos tradicionālā pieeja daudzuma, izmaksu un laika aplēsēm sāka strauji mainīties. Mūsdienu digitālās tehnoloģijas ļauj pilnībā automatizēt galvenos laika un izmaksu aprēķinus, nodrošinot lielāku precizitāti, ātrumu un pārredzamību būvniecības projektu resursu plānošanā.



Attēls 5.1-1 No daudzajām dažādajām sistēmām uzņēmējdarbībā vissvarīgākie ir rīki, kas atbild par apjoma, izmaksu un laika rādītājiem.

Būvniecības uzņēmumi koncentrējas uz precīziem laika un izmaksu datiem. Tie savukārt ir atkarīgi no izmantoto materiālu un darbaspēka apjoma, un to pārredzamība ietekmē rentabilitāti. Tomēr aprēķinu procesu sarežģītība un to nepārredzamība bieži vien noved pie lielākām projekta izmaksām, termiņu neievērošanas un pat bankrota.

Saskaņā ar KPMG ziņojumu "Pazīstamas problēmas - jaunas pieejas" (2023) tikai 50% būvniecības projektu tiek pabeigtī laikā, un 87% uzņēmumu ziņo par pastiprinātu kontroli pār kapitālprojektu ekonomiku. Galvenās problēmas ir saistītas ar kvalificēta personāla trūkumu un riska prognozēšanas grūtībām [2].

Vēsturiskie izmaksu aprēķina un procesu laika dati tiek apkopoti iepriekšējo projektu būvniecības laikā visā būvniecības uzņēmuma darbības laikā un ievadīti dažādu sistēmu datubāzēs (ERP, PMIS BPM, EPM u. c.).

Kvalitatīvi vēsturiskie izmaksu aprēķina dati ir būtiska būvniecības organizācijas konkurencē priekšrocība, kas tieši ietekmē tās izdzīvošanas spēju.

Būvniecības un inženiertehniskajos uzņēmumos ir izveidotas tāmēšanas un izmaksu aprēķinu nodajas, lai apkopotu, uzglabātu un atjauninātu vēsturiskos datus par projektu aprēķiniem. To galvenā funkcija ir uzkrāt un sistematizēt uzņēmuma pieredzi, kas laika gaitā palīdz uzlabot jaunu projektu apjoma, termiņu un izmaksu aplēšu precīzitāti. Šāda pieeja palīdz līdz minimumam samazināt kļūdas turpmākajos aprēķinos, pamatojoties uz jau īstenoto projektu praksi un rezultātiem.

## Projektu paredzamo izmaksu aprēkināšanas metodes

Izmaksu aprēķinu speciālisti izmanto dažādas aplēšu metodes, un katra no tām ir vērsta uz konkrētu datu veidu, informācijas pieejamību un projekta detalizācijas pakāpi. Visizplatītākās ir šādas:

- **Uz resursiem balstīta metode:** projekta izmaksu aplēšu aprēkināšana, pamatojoties uz visu nepieciešamo resursu, piemēram, materiālu, iekārtu un darbaspēka, detalizētu analīzi. Šai metodei nepieciešams detalizēts visu uzdevumu un katras uzdevuma veikšanai nepieciešamo resursu saraksts, kam seko to izmaksu aprēķins. Šī metode ir ļoti precīza, un to plaši izmanto izmaksu aplēsēs.
- **Parametriskā metode:** izmanto statistiskos modeļus, lai novērtētu izmaksas, pamatojoties uz projekta parametriem. Tas var ietvert izmaksu analīzi par mērvienību, piemēram, ēkas platību vai darbu apjomu, un šo vērtību pielāgošanu konkrētiem projekta nosacījumiem. Šī metode ir īpaši efektīva agrīnās stadijās, kad vēl nav pieejama detalizēta informācija.
- **Vienības metode (vienības izmaksu metode):** aprēķina projekta paredzamās izmaksas, pamatojoties uz izmaksām par mērvienību (piemēram, par kvadrātmetru vai kubikmetru). Tas nodrošina ātru un vienkāršu veidu, kā salīdzināt un analizēt dažādu projektu vai projektu daļu izmaksas.
- **Ekspertu vērtējums (Delfi metode):** balstās uz ekspertu viedokļiem, kuri izmanto savu pieredzi un zināšanas, lai novērtētu projekta vērtību. Šī pieeja ir noderīga, ja nav pieejami precīzi bāzes dati vai projekts ir unikāls.

Jāatzīmē, ka parametrisko metodi un ekspertu vērtējumus var pielāgot mašīnmācīšanās modeļiem. Tas ļauj automātiski ģenerēt projekta izmaksu un grafika prognozes, pamatojoties uz mācību paraugiem. Šādu modeļu izmantošanas piemēri sīkāk aplūkoti nodaļā "Mašīnmācīšanās piemērs projekta izmaksu un laika grafiku noteikšanai" ( 9.3-5. attēls).

Tomēr resursu metode joprojām ir vispopulārākā un visplašāk izmantotā metode pasaules praksē. Tā nodrošina ne tikai precīzu aplēsto izmaksu novērtējumu, bet arī ļauj aprēķināt atsevišķu procesu ilgumu būvlaukumā un visā projektā kopumā (vairāk informācijas nodaļā "Būvdarbu grafiki un 4D - projekta dati").

## **Uz resursiem balstīta metode izmaksu aprēķināšana un novērtēšana būvniecībā**

Uz resursiem balstīta izmaksu aprēķināšana ir vadības grāmatvedības metode, kurā projekta izmaksas tiek aprēķinātas, pamatojoties uz visu iesaistīto resursu tiešu uzskaiti. Būvniecībā šī pieeja ietver visu darbu veikšanai nepieciešamo materiālo, darbaspēka un tehnisko resursu detalizētu analīzi un novērtējumu.

Uz resursiem balstītā metode, nodrošina augstu pārredzamības un precizitātes pakāpi budžeta plānošanā, jo tā ir vērsta uz resursu faktiskajām cenām aplēšu veikšanas laikā. Tas ir īpaši svarīgi nestabilajā ekonomiskajā vidē, kur cenu svārstības var būtiski ietekmēt projekta kopējās izmaksas.

Turpmākajās nodaļās mēs detalizēti aplūkosim uz resursiem balstītu izmaksu aprēķināšanas procesu. Lai labāk izprastu tā principus būvniecībā, piemērosim analogiju ar vakariņu izmaksu aprēķinu restorānā. Restorāna vadītājs, plānojot vakaru, sastāda nepieciešamo produktu sarakstu, nēm vērā katru ēdienu pagatavošanas laiku un tad reizina izmaksas ar viesu skaitu. Būvniecībā process ir līdzīgs: katrai projekta elementu (objektu) kategorijai tiek veidotas detalizētas tāmes Receptes, un projekta kopējās izmaksas tiek noteiktas, summējot visas izmaksas kopējā rēķinā - galīgajā tāmē pa kategorijām.

Resursos balstītas pieejas galvenais un sākotnējais posms ir uzņēmuma sākotnējās datubāzes izveide. Pirmajā izmaksu aprēķina posmā tiek sastādīts strukturēts saraksts ar visiem priekšmetiem, materiāliem, darbu veidiem un resursiem, kas uzņēmuma rīcībā ir būvniecības projektu ietvaros - sākot ar naglu noliktavā un beidzot ar cilvēku aprakstu, norādot viņu kvalifikāciju un stundas likmi. Šī informācija tiek sistematizēta vienotā "Būvniecības resursu un materiālu datubāzē" - tabulārajā reģistrā, kas satur datus par nosaukumiem, raksturlielumiem, mērvienībām un pašreizējām cenām. Tieši šī datu bāze klūst par galveno un primāro informācijas avotu visiem turpmākajiem resursu aprēķiniem - gan izmaksu, gan darbu izpildes laika.

## **Būvniecības resursu datubāze: būvmateriālu un darbu katalogs**

Būvniecības resursu un materiālu datubāze vai tabula - ietver detalizētu informāciju par katru būvprojektā izmantojamo elementu - izstrādājumu, priekšmetu, materiālu vai pakalpojumu, tostarp tā

nosaukumu, aprakstu, mērvienību un vienības izmaksas, kas reģistrētas strukturētā veidā. Šajā tabulā var atrast visu, sākot no dažādiem projektos izmantotajiem kurināmā veidiem un materiāliem un beidzot ar detalizētiem speciālistu sarakstiem dažādu kategoriju veidā ar stundu likmju apraksti (5.1-2. att.).

## Database of resources

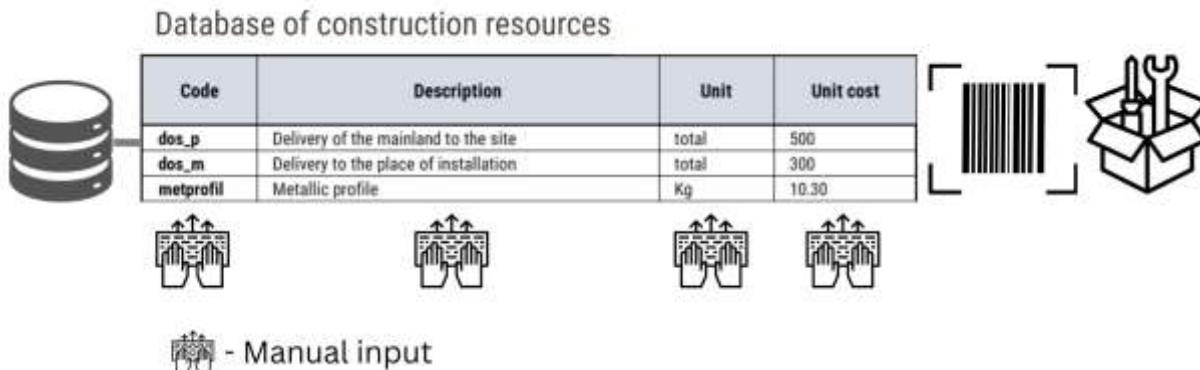
	1st grade potatoes	Sand lime bricks
<b>1 kg   \$2,99</b>	<b>1 pcs   \$1</b>	
	Black Angus marble beef	JCB 3CX backhoe loader
<b>1 kg   \$26,99</b>	<b>1 h   \$150</b>	
	Broccoli	Laborer of the 1st category
<b>1 pcs   \$1,99</b>	<b>1 h   \$30</b>	

Attēls 5.1-2 Resursu tabula ir sastāvdaļu saraksts, kas apraksta materiālu un pakalpojumu ar vienības izmaksām.

"Resursu datubāze" ir līdzīga tiešsaistes veikala preču katalogam, kurā katrai precei ir detalizēts tās īpašību apraksts. Tādējādi izmaksu aprēķinātājiem ir vieglāk izvēlēties pareizos resursus (piemēram, izvēloties produktus, pievienojot tos iepirkumu grozam), kas nepieciešami, lai aprēķinātu konkrētus būvniecības procesus aprēķinu veidā (galīgais pasūtījums tiešsaistes veikalā).

Resursu datubāzi var uzskatīt arī par visu restorāna pavārgrāmatas sastāvdaļu sarakstu. Katrs būvmateriāls, aprīkojums un pakalpojums ir līdzīgs receptēs izmantotajām sastāvdalām. "Resursu datubāze" ir detalizēts visu sastāvdaļu - būvmateriālu un pakalpojumu - saraksts, tostarp to izmaksas par vienību: gabalu, metru, stundu, litru utt.

Tabulai "Būvniecības resursu datubāzes" jaunus vienību elementus var pievienot divos veidos - manuāli ( 5.1-3. attēls) vai automātiski, integrējot ar uzņēmuma krājumu pārvaldības sistēmām vai piegādātāju datubāzēm.



**Attēls 5.1-3 Resursu datubāze tiek aizpildīta manuāli vai automātiski pārņem datus no citām datubāzēm.**

Tipisks vidēja lieluma būvniecības uzņēmums izmanto datu bāzi, kurā ir tūkstošiem un dažkārt pat desmitiem tūkstošu vienību ar detalizētiem aprakstiem, ko var izmantot būvniecības projektos. Šie dati tiek automātiski izmantoti līgumos un projekta dokumentācijā, lai precīzi aprakstītu darbu un procesu sastāvu.

Lai sekotu līdzī mainīgajiem tirgus apstākļiem, piemēram, inflācijai, katra produkta (preces vai pakalpojuma) atribūts "vienības izmaksas" resursu datubāzē ( 5.1-3. attēls) tiek regulāri atjaunināts manuāli vai automātiski lejupielādējot aktuālās cenas no citām sistēmām vai tiešsaistes platformām.

Resursu vienības izmaksu atjaunināšanu var veikt reizi mēnesī, ceturksnī vai gadā - atkarībā no resursa veida, inflācijas un ārējās ekonomiskās situācijas. Šādi atjauninājumi ir nepieciešami, lai saglabātu aprēķinu un aplēšu precīzitāti, jo šie pamatelementi ir izmaksu aprēķinātāju darba sākumpunkts. Jaunākos datus izmanto, lai sagatavotu tāmes, budžetus un grafikus, kas atspoguļo reālos tirgus apstākļus un samazina kļudu risku turpmākajos projekta aprēķinos.

### **Aprēķinu sastādīšana un darbu izmaksu aprēķināšana, pamatojoties uz resursu bāzi.**

Pēc tam, kad "Būvniecības resursu datu bāze" ( 5.1-3. attēls) ir piepildīta ar minimālajām vienību vienībām, varat sākt veidot aprēķinus, kas tiek aprēķināti katram procesam vai darbam būvlaukumā noteiktām mērvienībām: piemēram, vienam kubikmetram betona, vienam kvadrātmetram ģipškartona sienas, vienam metram apmales apmales vai vienam logu uzstādīšanas metram.

Piemēram, lai uzbūvētu 1 m<sup>2</sup> lielu kieģeļu sienu (attēls 5.1-4), pamatojoties uz iepriekšējo projektu pieredzi, ir nepieciešami aptuveni 65 kieģeļi (vienība "Silikāta kieģelis"), kuru cena ir 1 ASV dolārs par gabalu (atribūts "Cost per piece"), kopā 65 ASV dolāri. Turklat, pēc manas pieredzes, ir nepieciešams 10 minūtes izmantot celtniecības tehniku (vienība "JCB 3CX Loader"), kas kieģeļus novieto darba zonas tuvumā. Tā kā iekārtas noma maksā 150 ASV dolāru par stundu, 6 minūšu izmantošana izmaksātu aptuveni 15 ASV dolārus. Turklat 2 stundas būs nepieciešams kieģeļu ieklāšanas

darbuzņemējs, kura stundas likme ir 30 USD, bet kopējā summa ir 60 USD.

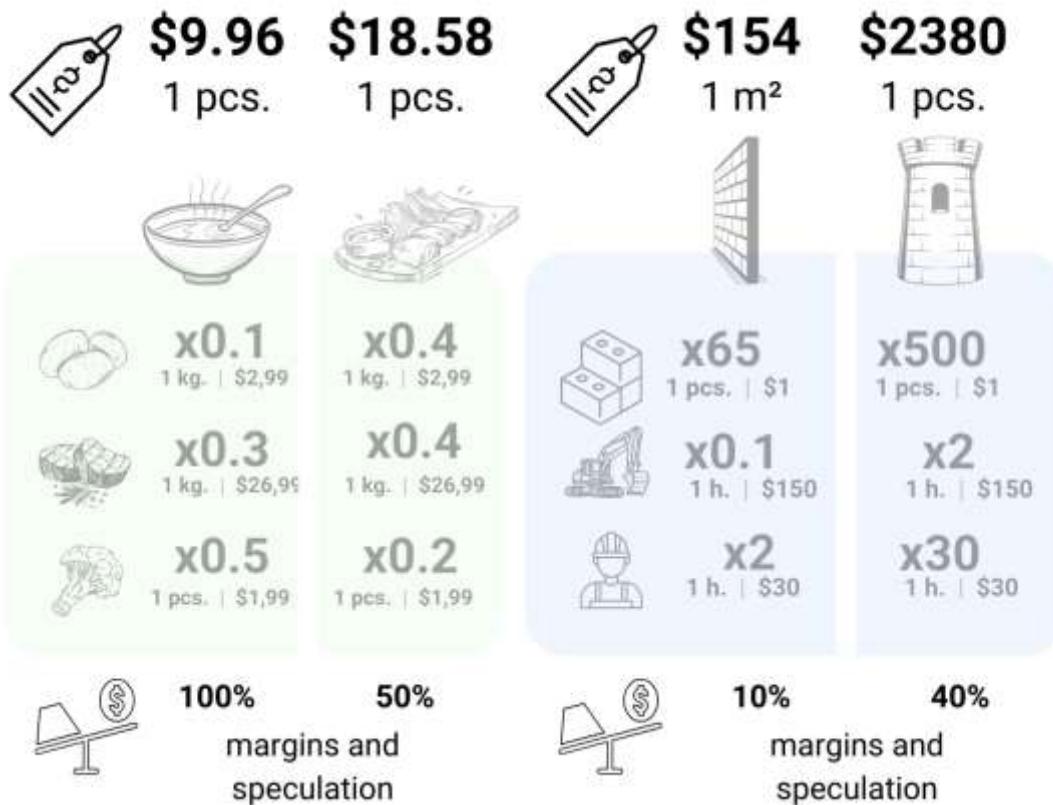
<b>x0.1</b> 1 kg   \$2,99	<b>x0.4</b> 1 kg.   \$2,99	<b>x65</b> 1 pcs   \$1	<b>x500</b> 1 pcs   \$1
<b>x0.3</b> 1 kg   \$26,99	<b>x0.4</b> 1 kg.   \$26,99	<b>x0.1</b> 1 h   \$150	<b>x2</b> 1 h   \$150
<b>x0.5</b> 1 pcs   \$1,99	<b>x0.2</b> 1 pcs.   \$1,99	<b>x2</b> 1 h   \$30	<b>x30</b> 1 h   \$30
	+ margins and speculation		+ margins and speculation
	+ margins and speculation		+ margins and speculation

**Attēlā: 5.1-4 Izmaksu aprēķinos sniepts detalizēts būvmateriālu un pakalpojumu saraksts, kas nepieciešami darbu un procesu veikšanai.**

Aprēķinu sastāvs (tā sauktās "receptes") tiek veidots, pamatojoties uz vēsturisko pieredzi, ko uzņēmums uzkrājis, veicot lielu daudzumu līdzīgu darbu. Šī praktiskā pieredze parasti tiek uzkrāta, izmantojot atsauksmes no būvlaukuma. Proti, meistars vāc informāciju tieši būvlaukumā, reģistrējot faktiskās darbaspēka izmaksas, materiālu patēriņu un tehnoloģisko operāciju nianses. Sadarbībā ar tāmēšanas nodalju šī informācija tiek iteratīvi precizēta: tiek precizēti procesu apraksti, koriģēts resursu sadalījums un izmaksu aprēķini tiek atjaunināti, lai atspoguļotu faktiskos datus, kas iegūti nesen īstenotajos projektos.

Tāpat kā receptē ir aprakstītas sastāvdaļas un daudzumi, kas nepieciešami ēdienu pagatavošanai, tā izmaksu tāmē ir detalizēts saraksts ar visiem būvmateriāliem, resursiem un pakalpojumiem, kas nepieciešami, lai pabeigtu konkrētu darbu vai procesu.

Regulāri veiktie darbi ļauj strādniekiem, meistariem un tāmētājiem orientēties nepieciešamajā resursu daudzumā: materiālos, degvielā, darba laikā un citos parametros, kas nepieciešami, lai veiktu darba vienību ( 5.1-5. attēls). Šie dati tiek ievadīti tāmēšanas sistēmās tabulu veidā, kur katrs uzdevums un operācija ir aprakstīta, izmantojot minimālos resursu bāzes elementus (ar pastāvīgi atjauninātām cenām), kas nodrošina aprēķinu precizitāti.



Attēls 5.1-5 Vienības likmes tiek apkopotas katram darbam, kur vienības apjoma atribūts tiek reizināts ar tās daudzumu, pievienojot peļņas procentu.

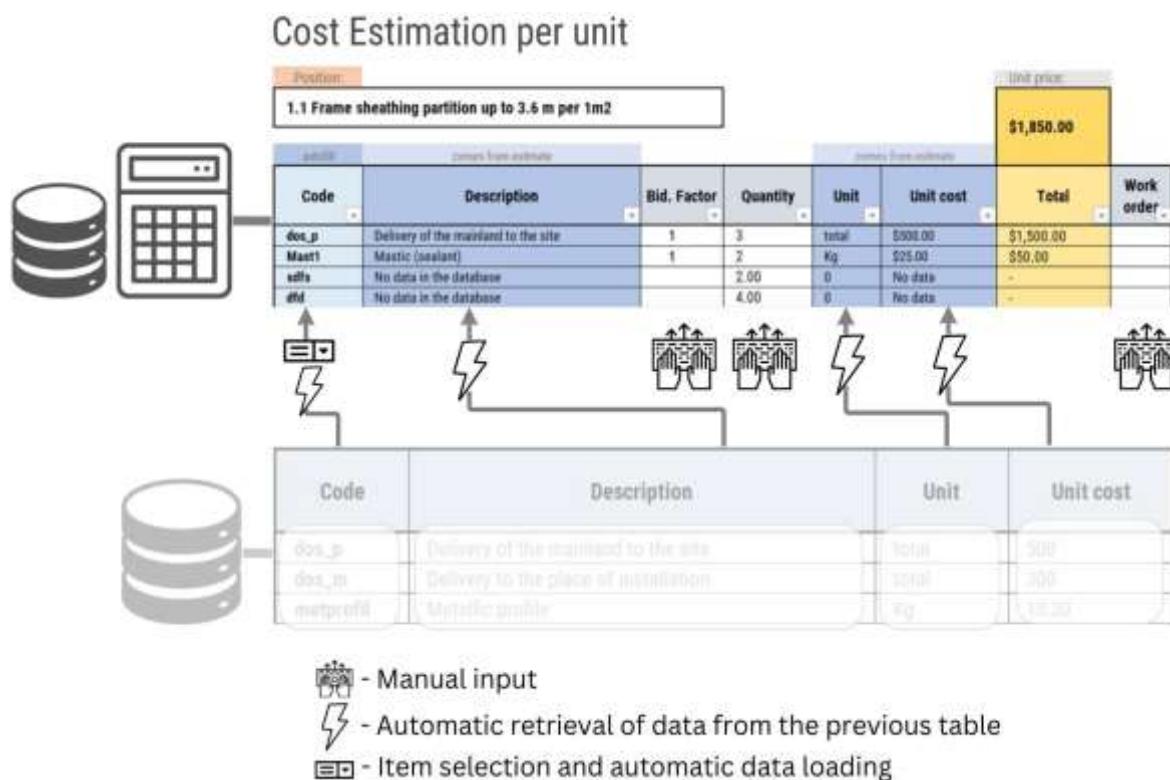
Lai iegūtu katra procesa vai darbības (pašizmaksas objekta) kopējās izmaksas, izmaksu raksturlielumu reizina ar tā skaitu un koeficientiem. Koeficientos var ķemt vērā dažādus faktorus, piemēram, darba sarežģītību, reģionālās īpatnības, inflācijas līmeni, iespējamos riskus (paredzamo pieskaitāmo izmaksu procentuālā daļa) vai spekulācijas (papildu peļņas koeficients).

Novērtētājs kā analītikis pārvērš meistara pieredzi un ieteikumus standartizētās tāmēs, aprakstot būvniecības procesus ar resursu vienību palīdzību tabulas formā. Būtībā tāmētāja uzdevums ir ar parametru un koeficientu palīdzību apkopot un strukturēt informāciju, kas nāk no būvlaukuma.

Tādējādi galīgās izmaksas par vienu darba vienību (piemēram, kvadrātmetru vai kubikmetru, vai vienas vienības uzstādīšanu) ietver ne tikai tiešās materiālu un darba izmaksas, bet arī uzņēmuma uzcenojumus, pieskaitāmās izmaksas, apdrošināšanu un citus faktorus ( 5.1-6. attēls)

Tajā pašā laikā mums vairs nav jāuztraucas par faktiskajām cenām (recepšu) aprēķinos, jo faktiskās cenas vienmēr ir atspoguļotas "resursu bāzē" (sastāvdaļu tabulā). Aprēķinu līmenī dati no resursu bāzes tiek automātiski ielādēti tabulā (piemēram, pēc preces koda vai tās unikālā identifikatora), kurā ielādēts apraksts un faktiskā cena par vienību, ko savukārt var automātiski ielādēt no tiešsaistes

platformām vai būvmateriālu tiešsaistes veikala. Aprēķinu līmenī darbu tāmētājam ir tikai jāapraksta darbs vai process, izmantojot atribūtu "resursu daudzums" un papildu faktorus.



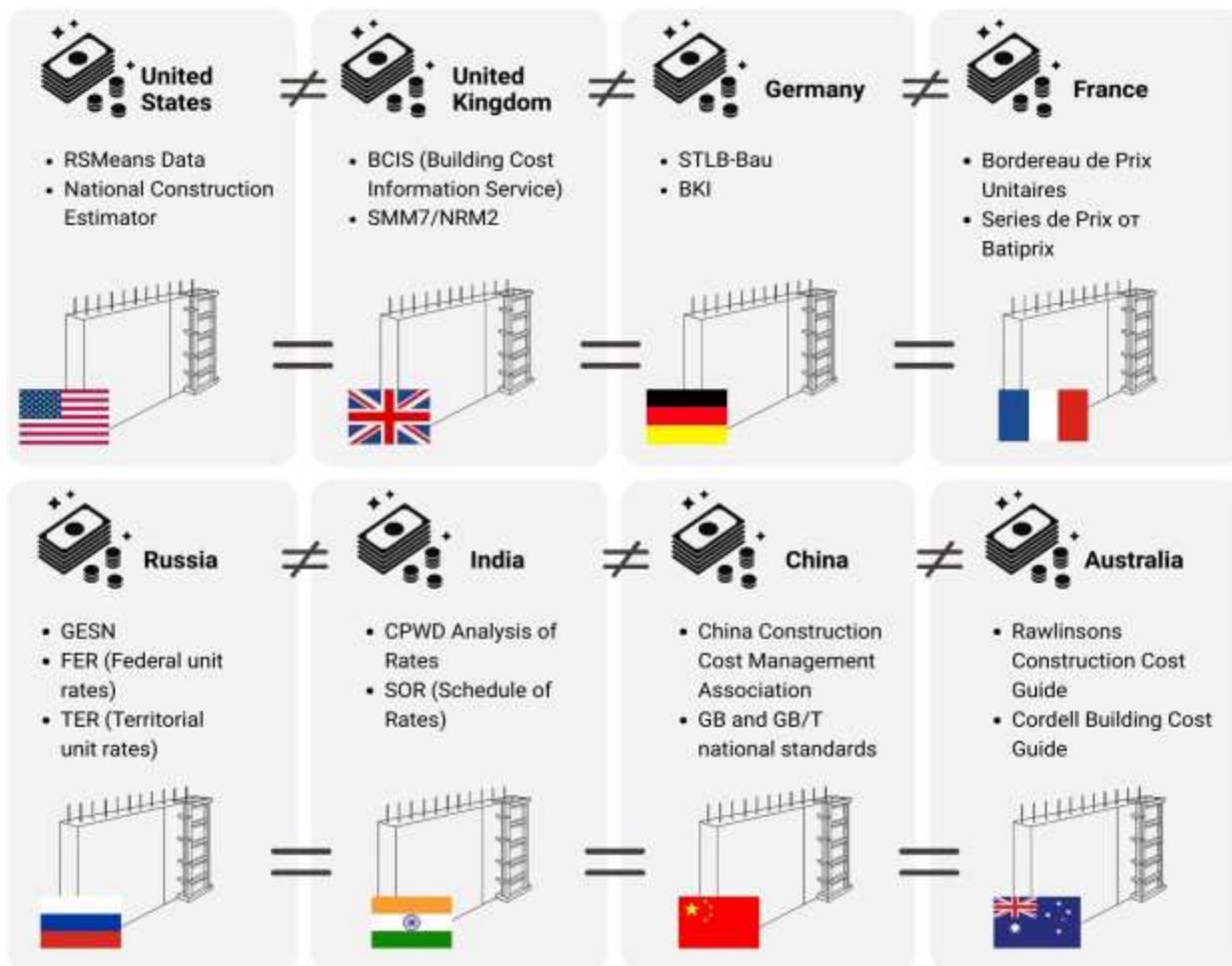
**Attēls 5.1-6 Darbu vienības izmaksu aprēķināšanas posmā tiek aizpildīti tikai nepieciešamo resursu daudzuma atribūti, viss pārējais tiek automātiski ielādēts no resursu datu bāzes.**

Izveidotās darba izmaksu kalkulācijas tiek uzglabātas tipisku projektu veidņu tabulās, kas ir tieši saistītas ar būvniecības resursu un materiālu datu bāzi. Šie veidnes paraugi ir standartizētas atkārtojošos darbu veidu receptes turpmākajiem projektiem, nodrošinot vienveidīgus aprēķinus visā uzņēmumā.

Kad datu bāzē mainās kāda resursa izmaksas ( 5.1-3. attēls) - vai nu manuāli, vai automātiski, lejupielādējot pašreizējās tirgus cenas (piemēram, inflācijas apstākļos), - atjauninājumi nekavējoties tiek atspoguļoti visos saistītajos izmaksu aprēķinos ( 5.1-6. attēls). Tas nozīmē, ka jāmaina tikai resursu bāze, bet izmaksu veidnes un tāmes laika gaitā paliek nemainīgas. Šāda pieeja nodrošina aprēķinu stabilitāti un atkārtojamību jebkādām cenu svārstībām, kas tiek ķemtas vērā tikai salīdzinoši vienkāršā resursu tabulā ( 5.1-3. attēls).

Katram jaunam projektam tiek izveidota standarta izmaksu veidnes kopija, kas ļauj veikt izmaiņas un pielāgot darbības, lai tās atbilstu konkrētām prasībām, nemainot uzņēmuma pieņemto sākotnējo veidni. Šāda pieeja nodrošina elastību aprēķinu pielāgošanā: jūs varat ķemt vērā būvlaukuma specifiku, klienta vēlmes, ieviest riska vai rentabilitātes (spekulācijas) koeficientus - tas viss, nepārkāpjot uzņēmuma standartus. Tas palīdz uzņēmumam atrast līdzsvaru starp peļņas maksimizēšanu, klientu apmierinātību un konkurētspējas saglabāšanu.

Dažās valstīs šādi izmaksu aprēķina paraugi, kas uzkrāti gadu desmitiem, ir standartizēti valsts līmenī un kļuvuši par daļu no valsts būvniecības izmaksu aprēķina sistēmas standartiem ( 5.1-7. attēls).



Attēlā: 5.1-7 Dažādās pasaules valstīs ir savi izmaksu aprēķināšanas noteikumi ar saviem (normatīvajiem) būvdarbu apkopojumiem un standartiem viena un tā paša elementa izmaksu aprēķināšanai.

Šādas standartizētas resursu aplēšu bāzes ( 5.1-7. attēls) ir obligāti jāizmanto visiem tirgus dalībniekiem jo īpaši attiecībā uz valsts finansētiem projektiem. Šāda standartizācija nodrošina

pārredzamību, salīdzināmību un taisnīgumu cenu veidošanā un līgumsaistību izpildē klientam

## **Projekta galīgo izmaksu aprēķins: no tāmes līdz budžetam**

Dažādās valstīs būvniecības praksē valsts un nozares aplēšu standarti ir atšķirīga nozīme. Lai gan dažās valstīs tiek prasīts stingri ievērot vienu standartu, lielākajā daļā attīstīto valstu ir elastīgāka pieeja. Tirgus ekonomikas valstīs valsts būvniecības standarti parasti kalpo tikai kā atskaites punkts. Būvniecības uzņēmumi šos standartus pielāgo saviem darbības modeļiem vai pilnībā tos pārstrādā, papildinot ar saviem pielāgotiem faktoriem. Šie pielāgojumi atspoguļo uzņēmumu pieredzi, resursu pārvaldības efektivitāti un bieži vien arī faktorus, kuros, piemēram, var īņemt vērā uzņēmuma spekulatīvo peļnu.

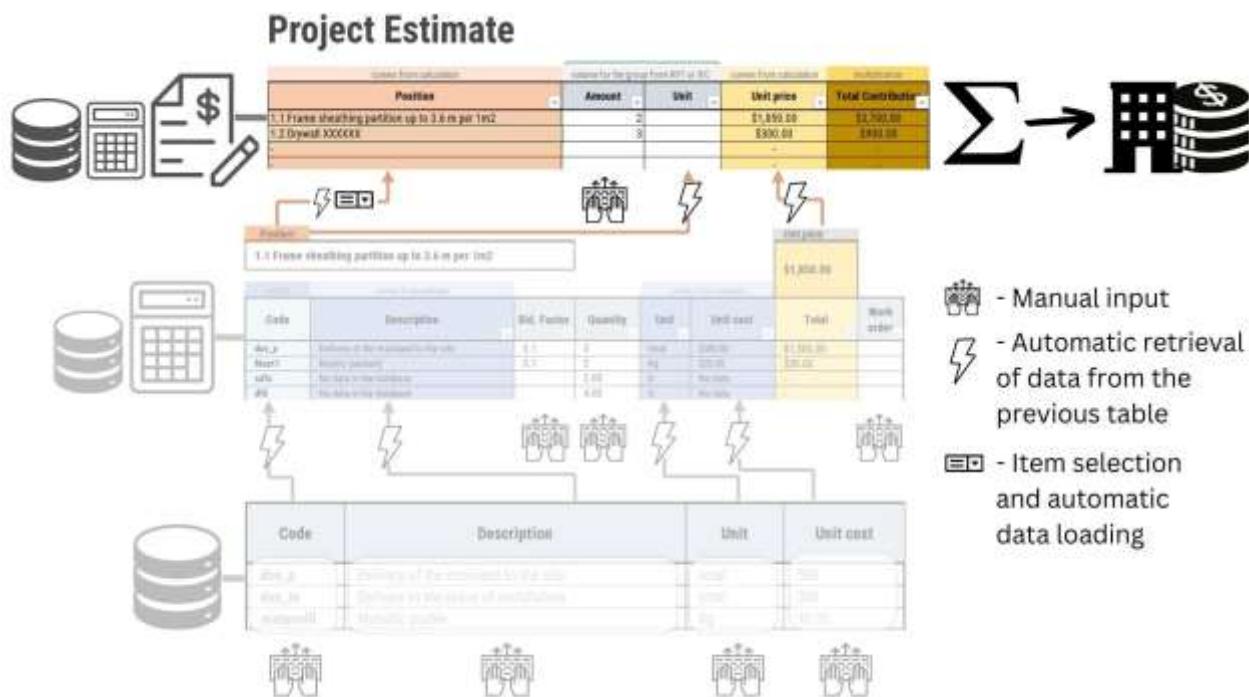
Rezultātā konkurences līmenis, tirgus pieprasījums, mērķa peļņa un pat attiecības ar konkrētiem klientiem var radīt būtiskas novirzes no standartizētajām normām. Šāda prakse nodrošina tirgus elastību, bet arī apgrūtina dažādu darbuzņēmēju piedāvājumu pārredzamu salīdzināšanu, ieviešot būvniecības nozarē šajā aprēķinu procesa posmā spekulatīvo cenu noteikšanas elementu.

Kad aprēķinu veidnes atsevišķām darbībām un procesiem ir sagatavotas - vai, biežāk, vienkārši nokopētas no valdības standarta tāmēm ( 5.1-7. attēls), pievienojot koeficientus, lai atspoguļotu konkrētā uzņēmuma "īpatnības" - , pēdējais solis ir katras posteņa izmaksas reizināt ar atbilstošo jaunā projekta darbu apjomu vai procesu atribūtu.

Aprēķinot jaunā būvniecības projekta kopējās izmaksas, galvenais solis ir apkopot visu izmaksu posteņu izmaksas, reizinot ar šo darbu apjomu projektā.

Lai izveidotu projekta kopējās izmaksas, mūsu vienkāršotajā piemērā mēs sākam ar viena kvadrātmetra sienas izbūves izmaksu aprēķinu un reizinām tās aprēķina izmaksas (piemēram, darbu "1m<sup>2</sup> standarta sienas elementu uzstādīšana") ar kopējo sienu kvadrātmetru skaitu projektā (piemēram, "Platība" vai "Daudzums" atribūts ( 5.1-8. attēls) no projekta CAD vai meistara aprēķini).

Izmaksas visiem projekta elementiem aprēķinām vienādi ( 5.1-8. attēls): īņemam izmaksas uz vienu darba vienību un reizinām tās ar konkrētā projekta elementa vai elementu grupas apjomu. Novērtētējam tikai jāievada šo elementu, darbību vai procesu skaits projektā apjoma vai daudzuma veidā. Tas ļauj automātiski ģenerēt pilnīgu būvniecības tāmi.



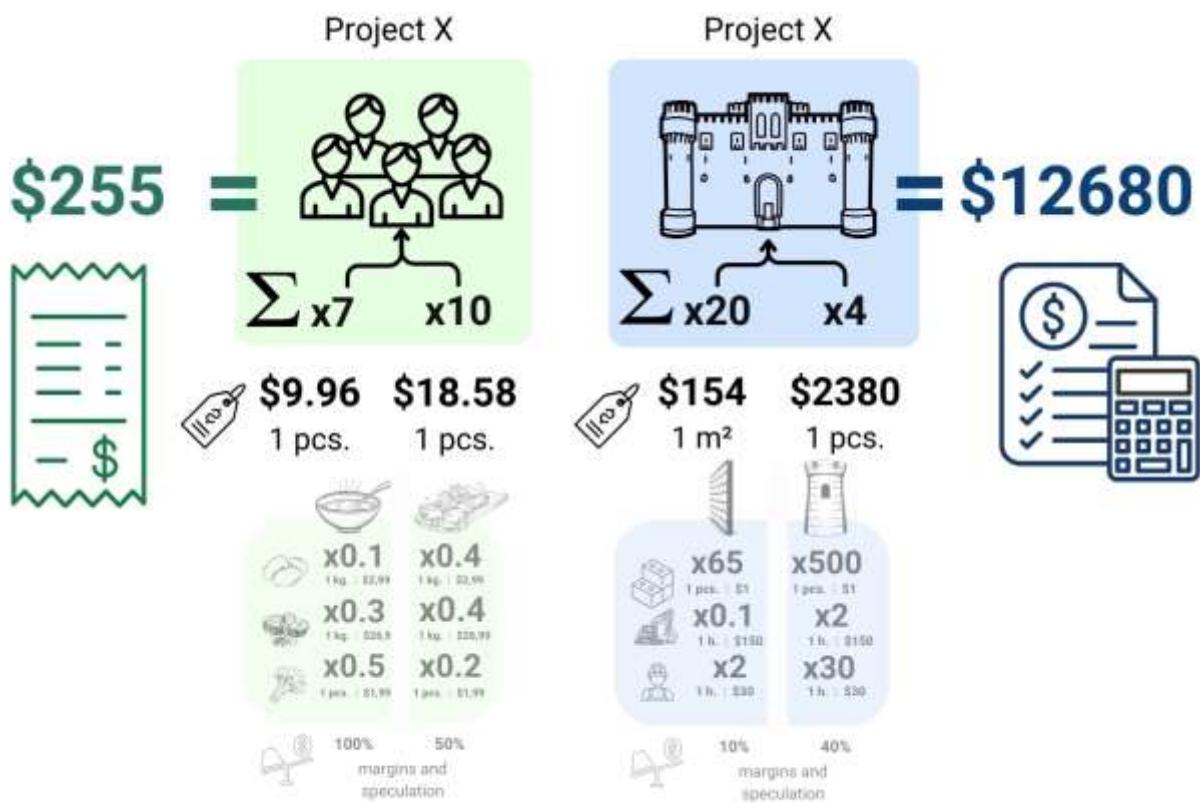
Attēls 5.1-8. Izmaksu tāmes izveides posmā mēs ievadām tikai darba apjomu.

Tāpat kā aprēķinu gadījumā, arī šajā līmenī mēs automātiski augšupielādējam gatavus aprēķinu posteņus (no aprēķinu veidnes vai jaunus, no veidnes kopētus un redīgētus), kuriem automātiski tiek pievienotas aktuālās izmaksas par darba vienību (kas tiek automātiski atjauninātas no resursu datu bāzes ( 5.1.1-8. attēls, apakšējā tabula)). Attiecīgi, ja resursu datubāzē vai izmaksu aprēķinu tabulās mainās dati - tāmē dati tieks automātiski atjaunināti aktuālajai dienai, nemainot pašus izmaksu aprēķinus vai pašu tāmi.

Restorāna kontekstā pasākuma galīgās izmaksas tiek aprēķinātas līdzīgi un ir vienādas ar visu vakariņu galīgajām izmaksām, kur katrā ēdienu izmaksas, reizinātas ar viesu skaitu, veido čeka kopējās izmaksas ( 5.1-9. attēls). Tāpat kā būvniecībā, arī ēdienu gatavošanas receptes restorānā var nemainīties gadu desmitiem. Atšķirībā no cenām, kur sastāvdaļu izmaksas var mainīties katru stundu.

Tāpat kā restorāna īpašnieks reizina katru ēdienu izmaksas ar porciju un cilvēku skaitu, lai noteiktu kopējās pasākuma izmaksas, tā arī izmaksu aprēķinu vadītājs saskaita visu projekta komponentu izmaksas, lai iegūtu pilnu būvniecības tāmi.

Tādējādi katrai aktivitātei projektā tiek noteiktas tās galīgās izmaksas ( 5.1-9. attēls), kas, reizinātas ar šai aktivitātei atbilstošās vienības atribūtu apjomu, dod aktivitāšu grupu izmaksas, no kurām iegūst visa projekta galīgās izmaksas.



Attēls 5.1-9 Galīgo tāmi aprēķina, summējot katras elementa darba izmaksu atribūtu ar tā darbības jomas atribūtu.

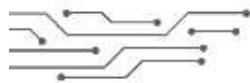
Projekta kopējās izmaksas ( 5.1.1-8. attēls) sniedz finansiālu priekšstatu par projektu, ļaujot klientiem, investoriem vai finanšu organizācijām izprast kopējo budžetu un finanšu resursus, kas nepieciešami projekta īstenošanai jebkurā konkrētajā dienā, nesmot vērā pašreizējās cenas.

Un, ja resursu bāzu, aprēķinu un tāmu sastādīšanas procesi (procesu receptes) jau ir izstrādāti, pusautomatizēti un slīpēti desmitiem tūkstošu gadu un reģistrēti valsts līmenī, tad automātiska kvalitatīvas informācijas iegūšana par elementu apjomu un daudzumu galīgās tāmes pēdējam posmam - šodien joprojām ir vāja vieta visu projekta izmaksu un laika atribūtu aprēķinu procesos un kopumā projekta kopējā budžeta veidošanā.

Tūkstošiem gadu tradicionālā tilpumu aprēķināšanas metode ir bijusi manuāla tilpuma un daudzuma mērīšana, izmantojot plakanu rasējumu. Līdz ar digitālā laikmeta iestāšanos uzņēmumi ir atklājuši, ka tagad informāciju par tilpumu un daudzumu var automātiski iegūt no ģeometriskiem datiem, kas ietverti CAD modeļos, tādējādi revolucionizējot tūkstošiem gadu vecos kvantitatīvo datu iegūšanas veidus.

Modernās pieejas procesu aplēšu un tāmu veidošanai ietver automātisku tilpuma un kvantitatīvo atribūtu iegūšanu no CAD datubāzēm, kuras var augšupielādēt un savienot ar izmaksu aprēķina procesu, lai iegūtu aktuālus projekta grupas apjomus jebkurā projektēšanas posmā līdz pat

ekspluatācijai.



## NODAĻA 5.2.

### **DAUDZUMA APRĒĶINS UN AUTOMĀTISKA TĀMJu UN GRAFIKU IZVEIDE**

#### **Pāreja no 3D uz 4D un 5D: volumetrisko un kvantitatīvo parametru izmantošana**

Kad rokās ir izmaksu aprēķina tabulas ar aprakstītajiem procesiem, izmantojot resursus ( 5.1-8. attēls), nākamais solis ir automātiski iegūt elementu grupas apjoma vai daudzuma parametrus, kas nepieciešami aprēķiniem un galīgajai tāmei.

Projekta elementu, piemēram, sienu vai plātņu, tilpuma rādītājus var automātiski iegūt no CAD datu bāzēm. CAD programmās izveidotos parametriskos objektus, izmantojot ģeometrijas kodolu, pārvērš garuma, platuma, laukuma, tilpuma un citu parametru skaitliskās vērtībās. Apjomu iegūšanas process, pamatojoties uz 3D ģeometriju, tiks sīkāk aplūkots nākamajā, sestajā daļā (6.3-3. att.), kas veltīta darbam ar CAD (BIM). Papildus apjomiem no CAD-modeļa datubāzes var iegūt arī līdzīgu elementu skaitu, filtrējot un grupējot objektus pēc kategorijām un īpašībām. Šie parametri, kas ļauj veikt grupēšanu, kļūst par pamatu projekta elementu sasaistei, izmantojot resursu aprēķinus, ar aprēķiniem, galīgo tāmi un visa projekta budžetu.

Tādējādi no 3D (CAD) modeļa iegūtais datu modelis tiek papildināts ar jauniem parametru slāniem, ko apzīmē kā 4D un 5D. Jaunajos vienības atribūtu slāņos, 4D (laiks) un 5D (izmaksas), 3D ģeometriskie dati tiek izmantoti kā vienības apjoma atribūtu vērtību avots.

- **4D** ir parametru informācijas slānis, kas elementu 3D parametriem pievieno informāciju par būvdarbu ilgumu. Šie dati ir būtiski, lai plānotu grafikus un pārvaldītu projekta termīpus
- **5D** ir nākamais datu modeļa paplašināšanas līmenis, kurā elementi tiek papildināti ar izmaksu raksturielumiem. Tādējādi ģeometriskajai informācijai tiek pievienots finansiālais aspekts - materiālu, darbu un iekārtu izmaksas, kas ļauj veikt budžeta aprēķinus, rentabilitātes analīzi un izmaksu pārvaldību būvniecības procesa laikā.

Izmaksas un 3D, 4D un 5D projekta vienību grupu atribūtu dati tiek aprakstīti līdzīgi aprēķiniem moduļu ERP, PIMS sistēmās (vai Excel līdzīgos rīkos) un tiek izmantoti automātiskai izmaksu aprēķināšanai un budžeta plānošanai gan atsevišķām grupām, gan visam projekta budžetam.

#### **5D atribūti un atribūtu apjomu iegūšana no CAD**

Sagatavojot būvprojekta galīgo tāmi, kuras sagatavošanu esam aplūkojuši iepriekšējās nodaļās (attēls 5.1-8), tilpuma atribūti katrai projekta elementu kategorijai tiek savākti vai nu manuāli, vai arī iegūti no tilpuma atribūtu specifikācijām, ko nodrošina CAD programmatūra.

Tradicionālā manuālā metode daudzumu aprēķināšanai paredz, ka meistars un tāmētājs analizē rasējumus, kas tūkstošiem gadu ir bijuši attēloti kā līnijas uz papīra un pēdējos 30 gadus - digitālā formātā, piemēram, PDF (PLT) vai DWG. Balstoties uz profesionālo pieredzi, viņi mēra nepieciešamo darbu un materiālu daudzumu, bieži vien izmantojot lineālu un transportieri. Šī metode prasa ievērojamas pūles un laiku, kā arī īpašu uzmanību detaļām.

Šādā veidā noteikt darbības jomas atribūtus var aizņemt no dažām dienām līdz vairākiem mēnešiem atkarībā no projekta apjoma. Turklāt, tā kā visi mērījumi un aprēķini tiek veikti manuāli, pastāv cilvēciskas kļūdas risks, kas var novest pie neprecīzu datu iegūšanas, kas pēc tam ietekmē kļūdas projekta laika un izmaksu aplēsēs, par kurām atbildīgs būs viss uzņēmums.

Mūsdienu metodes, kas balstītas uz CAD datu bāzu izmantošanu, ievērojami vienkāršo tilpumu aprēķināšanu. CAD modeļos elementu ģeometrija jau ietver tilpuma atribūtus, kurus var automātiski aprēķināt (izmantojot ģeometrijas kodolu ( 6.3-3. attēls)) un attēlot vai eksportēt tabulas formā.

Šādā scenārijā tāmēšanas nodaļa pieprasā CAD projektētājam datus par projekta elementu daudzuma un apjoma parametriem. Šie dati tiek eksportēti kā izklājlapas vai tieši integrēti izmaksu aprēķinu datubāzēs - Excel, ERP vai PMIS - sistēmās. Šis process bieži vien nesākas ar oficiālu pieprasījumu, bet gan arī dialogu starp klientu (iniciatoru) un arhitektu tāmētāju no būvniecības vai projektēšanas uzņēmuma. Zemāk ir sniegti vienkāršots piemērs, kurā parādīts, kā no ikdienas komunikācijas tiek veidota strukturēta tabula automātiskajiem aprēķiniem (QTO):

- ⌚ Klients - "Es vēlos piebūvēt ēkai vēl vienu stāvu tādā pašā konfigurācijā kā pirms stāvs".
- ⌚ Arhitekts (CAD) - "Pievienojot trešo stāvu, konfigurācija ir tāda pati kā otrajā stāvā". Un pēc šī ziņojuma projekta CAD versiju nosūta tāmētājam.
- ⌚ Aplēšu kalkulators automātiski veic grupēšanu un aprēķinus (ERP, PMIS, Excel) - "Es palaidīšu projektu caur Excel izklājlapu ar QTO noteikumiem (ERP, PMIS), iegūšu apjomus pa kategorijām jaunajam stāvam un izveidoju tāmi"

Rezultātā teksta dialogs tiek pārveidots par tabulas struktūru ar grupēšanas noteikumiem:

Elements	Kategorija	Grīda
Pārkāšanās	OST_Floors	3
Kolonna	OST_StructuralColumns	3
kāpnes	OST_Stairs	3

Pēc projektētāja CAD modeļa automātiskās grupēšanas procesa saskaņā ar QTO aprēķinu kalkulatora noteikumiem un automātiskās apjomu reizināšanas ar resursu aprēķiniem ( 5.1-8. attēls) tiek iegūti šādi rezultāti, kas tiek nosūtīti klientam:

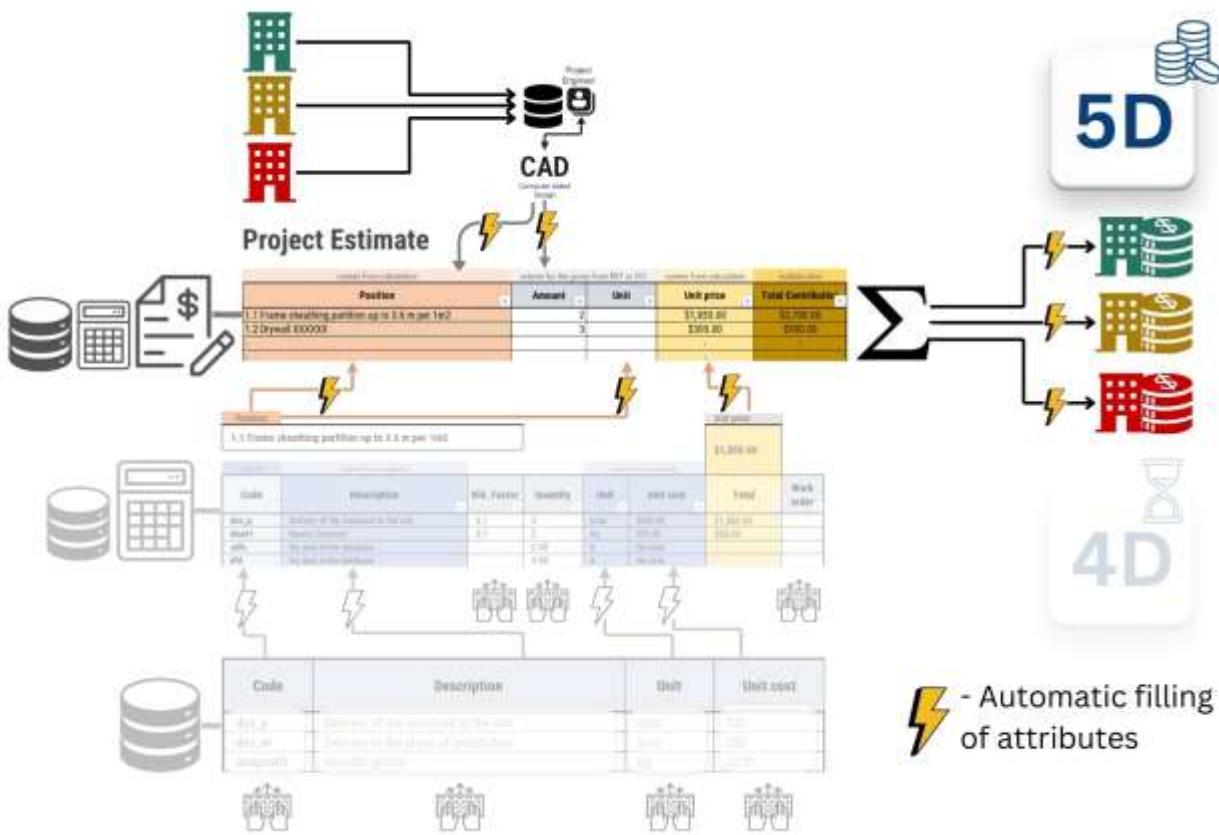
Elements	Tilpums	Grīda	Cena par vienību.	Kopējās izmaksas
Pārkāšanās	420 M <sup>2</sup>	3	150 €/M <sup>2</sup>	<b>63 000 €</b>
Kolonna	4 gab.	3	2450 € par gabalu.	<b>9 800 €</b>
kāpnes	2 gab.	3	4 300 € par gabalu.	<b>8 600 €</b>
<b>KOPĀ:</b>	-	-	-	<b>81 400 €</b>

- ⌚ Klients - "Paldies, tas ir diezgan daudz, mums vajag samazināt dažas telpas". Un cikls atkārtojas daudzas reizes.

Šis scenārijs var atkārtoties daudzas reizes, jo īpaši apstiprināšanas posmā, kad klients sagaida tūlītēju atgriezenisko saiti. Tomēr praksē šādi procesi var ieilgt dienām vai pat nedēļām. Mūsdienās, pateicoties automātiskās grupēšanas un aprēķinu noteikumu ieviešanai, darbības, kas agrāk prasīja daudz laika, ir jāveic dažu minūšu laikā. Automatizēta daudzumu iegūšana, izmantojot grupēšanas noteikumus, ne tikai paātrina aprēķinus un aplēses, bet, samazinot līdz minimumam cilvēcisko faktoru, samazina kļūdu iespējamību, nodrošinot pārrēķināmu un precīzu projekta izmaksu novērtējumu.

Ja, veidojot 3D modeli CAD sistēmā, sākotnēji tika ņemtas vērā novērtēšanas nodaļas prasības (kas praksē joprojām ir reti sastopams) un elementu grupu nosaukumi, identifikatori un to klasifikācijas atribūti ir noteikti parametru veidā, kas sakrīt ar novērtēšanas grupu un klašu struktūrām, tad tilpuma atribūtus var automātiski pārnest uz novērtēšanas sistēmām bez papildu transformācijām.

Automātiska tilpuma atribūtu ieguve no CAD specifikāciju tabulu veidā ļauj ātri iegūt aktuālus datus par atsevišķu darbu izmaksām un projektu kopumā ( 5.2-1. attēls). Atjauninot tikai CAD failu ar projekta apjomiem aprēķinu procesā vai aprēķinu sistēmā, uzņēmums var ātri pārrēķināt tāmi, ņemot vērā jaunākās izmaiņas, nodrošinot augstu visu turpmāko aprēķinu precizitāti un konsekvenci.



Attēls 5.2-1 Apjoma atribūti no CAD tabulām vai datubāzēm tiek automātiski ievadīti tāmē, jaujot uzreiz aprēķināt kopējās projekta izmaksas.

Pieaugot kapitālprojektu sarežģītībai, pilna budžeta aprēķināšana un projektu kopējo izmaksu analīze saskaņā ar šādu scenāriju ( 5.2-1. attēls ) - kļūst par galveno rīku pamatotu lēmumu pieņemšanai.

Saskaņā ar Accenture pētījumu Creating More Value through Capital Projects (2024) [20] vadošie uzņēmumi aktīvi integrē datu analītiku digitālajās iniciatīvās, izmantojot vēsturisko informāciju, lai prognozētu un optimizētu rezultātus. Pētījumi liecina, ka arvien vairāk īpašnieku un apsaimniekotāju izmanto lielo datu analītiku, lai prognozētu tirgus tendencies un novērtētu komerciālo dzīvotspēju pirms projektēšanas uzsākšanas. Tas tiek panākts, analizējot datu krātuves no esošā projektu portfeļa. Turklāt 79% īpašnieku-operatoru ievieš "spēcīgu" prognozēšanas analītiku, lai novērtētu projekta rezultātus un atbalstītu operatīvo lēmumu pieņemšanu reāllaikā.

Mūsdienu efektīva būvniecības projektu vadība ir nesaraujami saistīta ar liela apjoma informācijas apstrādi un analīzi visos projektēšanas posmos un procesos pirms projektēšanas. Izmantojot datu noliktavas, resursu aprēķinus, prognozēšanas modeļus un mašīnmācīšanos, iespējams ne tikai samazināt riskus aprēķinos, bet arī pieņemt stratēģiskus lēmumus par projekta finansēšanu agrīnās projektēšanas stadijās. Par datu noliktavām un prognozēšanas modeļiem, kas papildinās aprēķinus, plašāk runāsim grāmatas devītajā daļā.

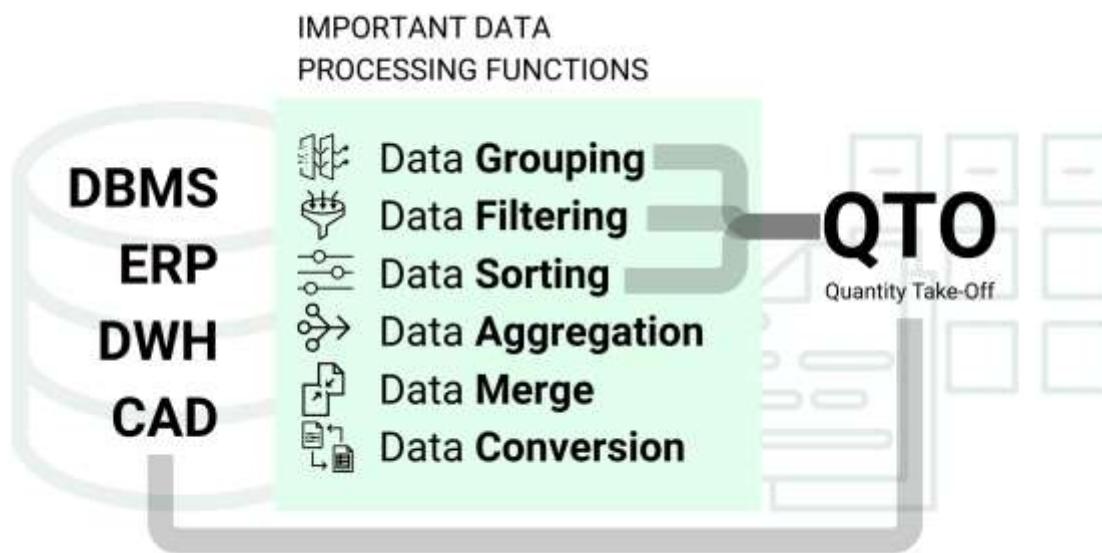
Automātiska elementu tilpuma parametru iegūšana no CAD projektiem, kas nepieciešami tāmju sagatavošanai, tiek veikta ar grupēšanas rīku QTO (Quantity Take-Off) palīdzību. QTO rīki darbojas,

grupējot visus projekta objektus pēc īpašiem elementu identifikatoriem vai elementu atribūtu parametriem, izmantojot CAD datubāzē izveidotās specifikācijas un tabulas.

## QTO Quantity Take-Off: projekta datu grupēšana pēc atribūtiem

QTO (Quantity Take-Off) būvniecībā ir projekta īstenošanai nepieciešamo elementu daudzuma noteikšanas process. Praksē QTO bieži vien ir daļēji manuāls process, kas ietver datu vākšanu no dažādiem avotiem: PDF dokumentiem, DWG rasējumiem un digitālajiem CAD modeļiem.

Strādājot ar datiem, kas iegūti no CAD datubāzēm, QTO process tiek īstenots kā filtrēšanas, šķirošanas, grupēšanas un apkopošanas darbību secība. Modeļa elementi tiek atlasīti atbilstoši klašu, kategoriju un tipu parametriem, pēc tam to kvantitatīvie atribūti, piemēram, tilpums, laukums, garums vai daudzums, tiek apkopoti atbilstoši aprēķinu logikai ( 5.2-2. attēls).



Attēls 5.2-2 Datu grupēšana un filtrēšana ir populārākās funkcijas, ko izmanto datu bāzēs un datu noliktavās.

QTO (filtrēšanas un grupēšanas) process ļauj sistematizēt datus, veidot specifikācijas un sagatavot ievades informāciju tāmju, iepirkumu un darba grafiku aprēķināšanai. QTO pamatā ir elementu klasifikācija pēc izmērāmo atribūtu veida. Katram elementam vai elementu grupai tiek izvēlēts atbilstošais kvantitatīvais mēriju parametrs. Piemēram:

- **Garuma atribūts** (apmales akmens - metros)
- **Platības raksturlielums** (gipškartona darbi - kvadrātmetros)
- **Tilpuma raksturlielums** (betona darbi - kubikmetros)
- **Daudzuma raksturlielums** (logi - par gabalu)

Papildus tilpuma raksturojumam, kas matemātiski izveidots, pamatojoties uz ģeometriju, aprēķinos pēc QTO grupēšanas bieži vien tiek piemēroti pārtēriņa koeficienti ( 5.2-12. attēls, piemēram, 1,1, lai

ņemtu vērā 10% loģistikas un uzstādīšanas izmaksām) - korekcijas vērtības, kas ņem vērā zudumus, uzstādīšanu, uzglabāšanu vai transportēšanu. Tas ļauj precīzāk prognozēt faktisko materiālu patēriņu un izvairīties gan no materiālu trūkuma, gan pārpalikuma būvlaukumā.

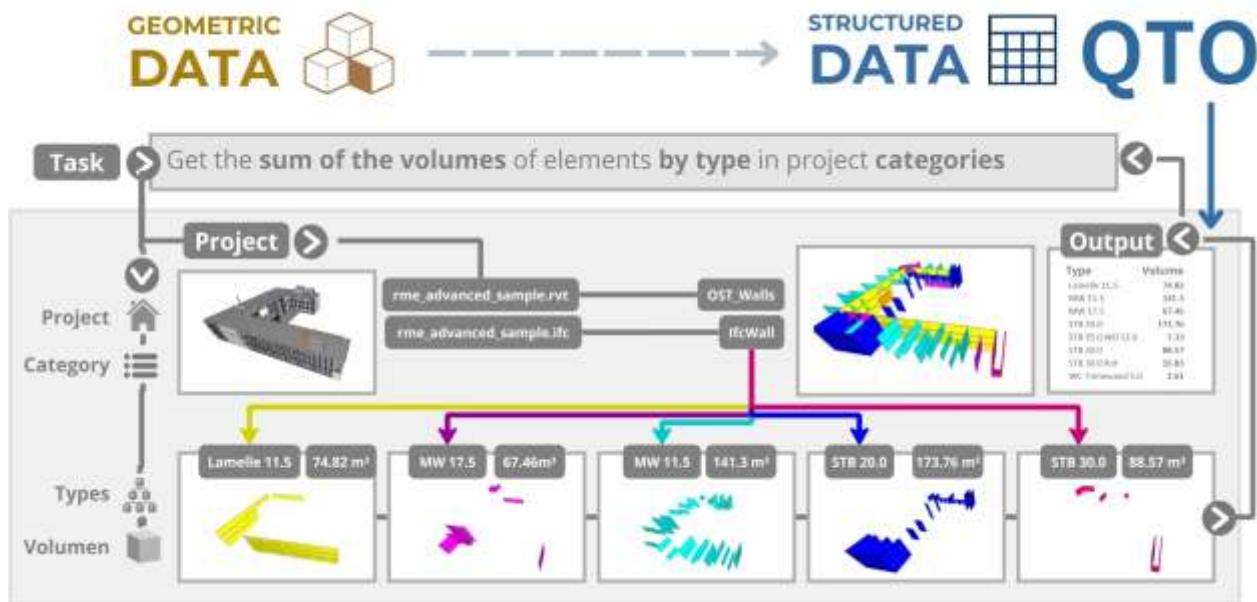
Automatizēts daudzuma ņemšanas process (QTO) ir būtisks, lai sagatavotu precīzus aprēķinus un aplēses, samazinātu cilvēcisko klūdu apjomu specifikācijas procesos un novērstu pārāk lielu vai nepietiekamu materiālu pasūtīšanu.

Kā piemēru QTO procesam aplūkosim parastu gadījumu, kad no CAD datu bāzes ir nepieciešams parādīt tilpumu tabulu-specifikāciju pēc elementu tipiem noteiktai elementu kategorijai, klasei. Sagrupēsim visus projekta elementus pēc tipa no CAD projekta sienas kategorijas un apkoposim katras tipa tilpuma atribūtus, lai rezultātu parādītu kā QTO tilpuma tabulu ( 5.2-3. attēls).

Tipiska CAD projekta piemērā ( 5.2-3. attēls) visi sienu kategorijas elementi CAD datubāzē ir sagrupēti pēc sienu tipa, piemēram, "Lamelle 11.5", "MW 11.5" un "STB 20.0", un tiem ir skaidri definēti tilpuma atribūti, kas attēloti metriskajos kubos.

Vadītāja, kas ir saskarne starp projektētājiem un aprēķinu speciālistiem, mērķis ir iegūt automatizētu tabulu ar tilpumiem pa elementu tipiem izvēlētajā kategorijā. Ne tikai konkrētam projektam, bet arī universālā formā, kas piemērojama citiem projektiem ar līdzīgu modeļa struktūru. Tas nodrošina pieejas mērogojamību un ļauj atkārtoti izmantot datus, nedublējot pūles.

Ir pagājuši tie laiki, kad pieredzējuši projektētāji un tāmētāji bija bruņojušies ar lineālu un rūpīgi mērīja katru līniju uz papīra vai PDF plānos - tradīcija, kas nav mainījusies pēdējo gadu tūkstošu laikā. Līdz ar 3D modelēšanas attīstību, kur katra elementa ģeometrija tagad ir tieši saistīta ar automātiski aprēķinātiem tilpuma atribūtiem, tilpumu un QTO daudzumu noteikšanas process ir kļuvis automatizēts.



Attēls 5.2-3 QTO apjoma un daudzuma atribūtu iegūšana no projekta ietver projekta elementu grupēšanu un filtrēšanu.

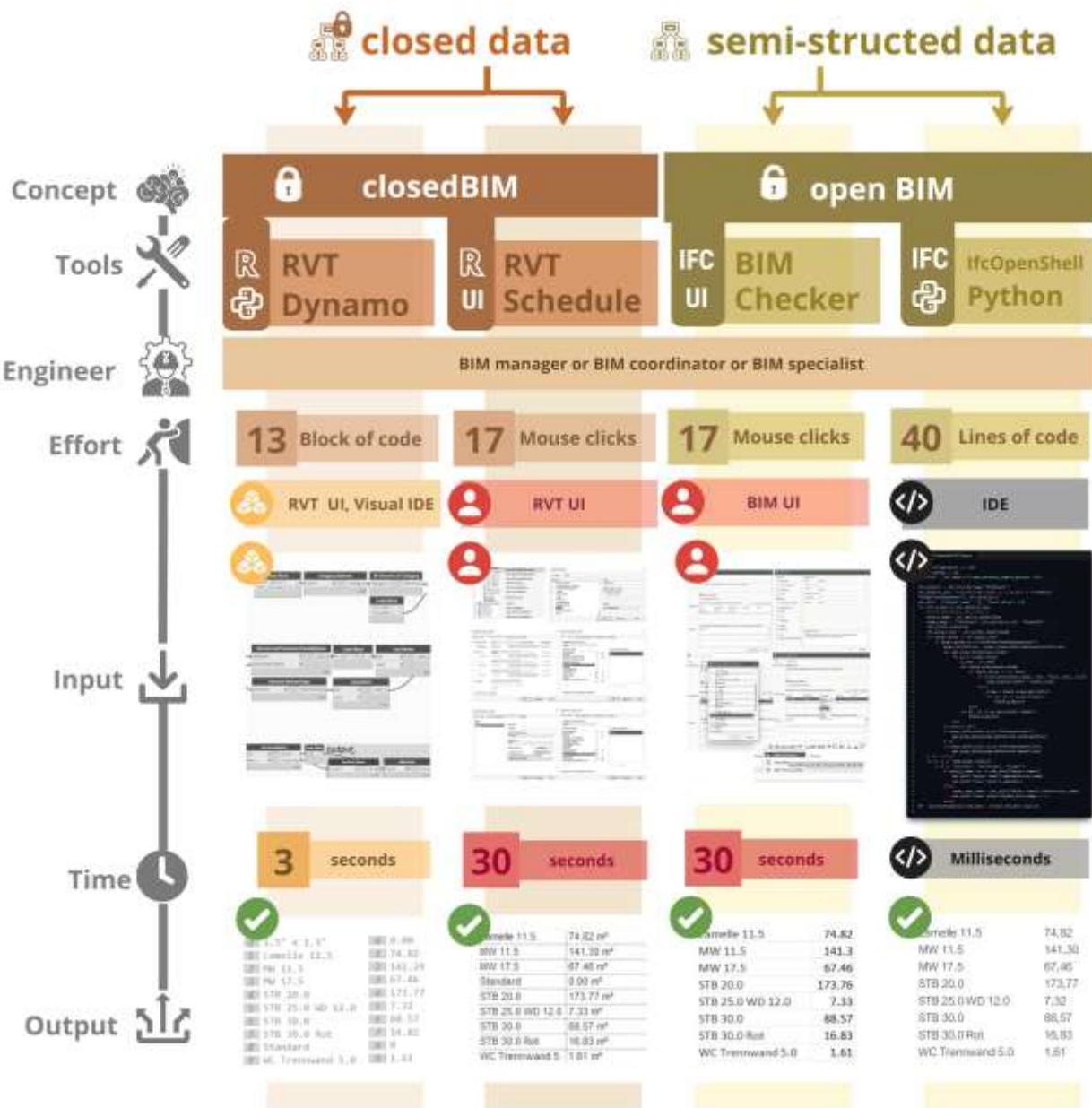
Mūsu piemērā uzdevums ir "atlaist sienu kategoriju projektā, sagrupēt visus elementus pēc tipa un strukturētā tabulas formātā sniegt informāciju par apjoma atrībūtiem", lai šo tabulu varētu izmantot desmitiem citu speciālistu izmaksu aprēķinu, logistikas, darbu grafiku un citu darījumu aprēķinu veikšanai ( 6.1-3. attēls).

Sakarā ar CAD datu slēgtību ne katrs speciālists šodien var izmantot tiešu piekļuvi CAD datu bāzei (piekļuves problēmas iemesli un risinājumi ir sīki izklāstīti grāmatas sestajā daļā). Tāpēc daudziem nākas izmantot specializētus BIM rīkus, kas balstīti uz atvērtā BIM un slēgtā BIM koncepcijām [63]. Strādājot ar specializētajiem BIM rīkiem vai tieši CAD programmatūras vidē, tabulu ar QTO (Quantity Take-Off) rezultātiem var izveidot dažādos veidos - atkarībā no tā, vai tiek izmantota manuāla saskarne vai programmatūras automatizācija.

Piemēram, izmantojot CAD (BIM) programmatūras lietotāja interfeisu, pietiek veikt aptuveni 17 darbības (pogu klikšķus), lai iegūtu gatavu tilpumu tabulu ( 5.2-4. attēls). Tomēr lietotājam ir labi jāpārzina CAD (BIM) programmatūras modeļa struktūra un funkcijas.

Ja automatizācija tiek veikta, izmantojot programmas kodu vai CAD programmu spraudņus un API rīkus, samazinās manuālo darbību skaits, lai iegūtu tilpuma tabulas, taču atkarībā no izmantotās bibliotēkas vai rīka būs jāraksta no 40 līdz 150 rindīgām koda:

- **IfcOpSh (atvērts BIM)** vai **Dynamo IronPython (slēgts BIM)** - ja už iegūt QTO tabulu no CAD formāta vai CAD programmas tikai ~40 kodu rindās.
- **IFC\_js (atvērts BIM)** - lai no IFC modeļa iegūtu apjomīgus atrībūtus, ir nepieciešamas aptuveni 150 rindas koda.
- **Saskarnes CAD rīki (BIM)** - ar 17 peles klikšķiem var iegūt tādu pašu rezultātu manuāli.



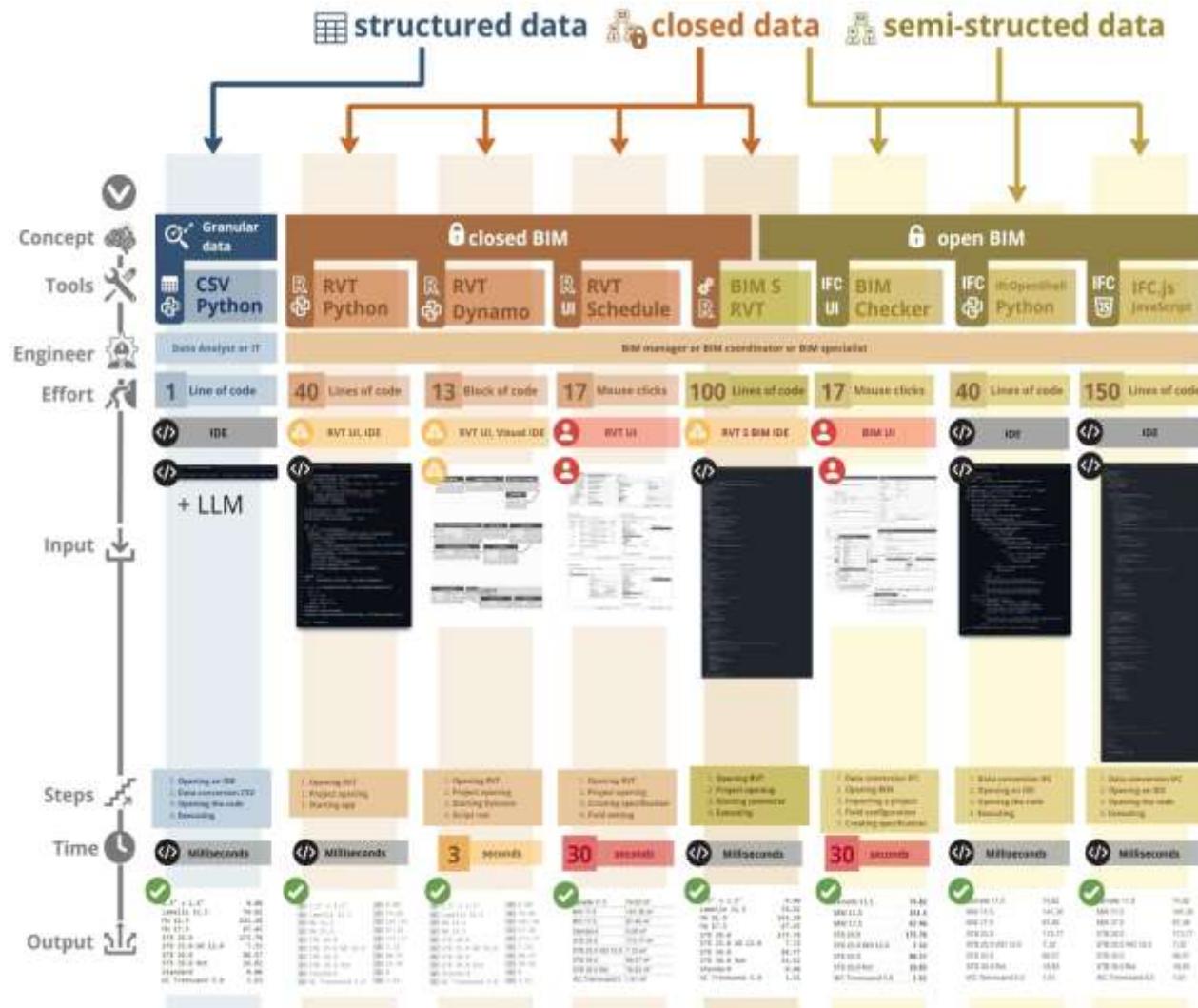
Attēls 5.2-4 CAD (BIM) projektētāji un vadītāji QTO tabulu izveidei izmanto no 40 līdz 150 rindiņām koda vai dučiem taustiņu nos piedumu.

Rezultāts ir tāds pats - strukturēta tabula ar elementu grupas apjoma atribūtiem. Vienīgā atšķirība ir darbaspēka izmaksas un lietotājam nepieciešamās tehniskās zināšanas ( 5.2-4. attēls). Mūsdienīgi rīki saīdzinājumā ar apjoma manuālu vākšanu ievērojami paātrina KTO procesu un samazina kļūdu iespējamību. Tie ļauj iegūt datus tieši no projekta modeļa, novēršot nepieciešamību manuāli pārrēķināt apjomus no rasējumiem, kā tas tika darīts agrāk.

Neatkarīgi no izmantotās metodes - vai tas ir atvērtais BIM vai slēgtais BIM - ir iespējams iegūt identisku QTO - tabulu ar projekta elementu apjomiem ( 5.2-4. attēls). Tomēr, strādājot ar projekta

datiem CAD - (BIM-) koncepcijās, lietotāji ir atkarīgi no specializētiem rīkiem un API, ko nodrošina ražotāji ( 3.2-13. att.). Tas rada papildu atkarības slāņus un prasa apgūt unikālas datu shēmas, vienlaikus ierobežojot tiešu piekļuvi datiem.

Sakarā ar CAD datu slēgto raksturu, QTO tabulu un citu parametru iegūšana sarežģī aprēķinu automatizāciju un integrāciju ar ārējām sistēmām. Izmantojot rīkus tiešai piekļuvei datubāzēm un ar reversās inženierijas rīku palīdzību pārnesot projekta CAD -datus atvērtā strukturētā datu rāmja formātā ( 4.1-13. attēls), identisku QTO tabulu var iegūt tikai ar vienu koda rindu ( 5.2-5. attēls - variants ar granulāriem datiem).



Attēls 5.2-5 Dažādi rīki sniedz vienādus rezultātus projekta vienību atribūtu tabulu veidā, bet ar dažādām darba izmaksām.

Izmantojot atvērtus strukturētus datus no CAD projektiem, kā minēts nodaļā "CAD (BIM) datu konvertēšana strukturētā formā", grupēšanas process, QTO, ir ievērojami vienkāršots.

Pieejas, kuru pamatā ir atvērtu strukturētu datu izmantošana vai tieša piekļuve CAD modeļu datubāzēm, ir brīvas no mārketinga ierobežojumiem, kas saistīti ar akronīmu BIM. Tās balstās uz

pārbaudītiem rīkiem, kas jau sen tiek izmantoti citās nozarēs ( 7.3-10. attēls ETL process).

Saskaņā ar McKinsey pētījumu "Open Data: Atdzīvināt inovācijas un produktivitāti ar straumētas informācijas paīdzību". [102], kas tika veikts 2013. gadā, atvērto datu izmantošana varētu radīt iespējas ietaupīt no 30 līdz 50 miljardiem ASV dolāru gadā elektroenerģijas ražošanas iekārtu projektēšanas, projektēšanas, iepirkuma un būvniecības jomā. Tas nozīmē 15% ietaupījumu būvniecības kapitālieguldījumu izmaksās.

Darbs ar atvērtiem strukturētiem (granulāriem) datiem vienkāršo informācijas iegūšanu un apstrādi, samazina atkarību no specializētām BIM platformām un paver ceļu automatizācijai bez nepieciešamības izmantot patentētas sistēmas vai parametriskus un sarežģītus CAD formātu datu modeļus.

## **QTO automatizācija, izmantojot LLM un strukturētus datus**

Nestrukturizētu datu tulkošana strukturētā formā ievērojami uzlabo dažādu procesu efektivitāti: tā vienkāršo datu apstrādi ( 4.1-1. attēls, 4.1-2. attēls) un paātrina validācijas procesu, padarot prasības skaidras un pārskatāmas, kā jau tika aplūkots iepriekšējās nodalās. Tāpat CAD datu (BIM) tulkošana strukturētā atvērtā formā ( 4.1-12. attēls, 4.1-13. attēls) atvieglo atribūtu grupēšanas procesu un QTO procesu.

QTO atribūtu tabulai ir strukturēta forma, tāpēc, izmantojot strukturētus CAD datus, mēs strādājam ar vienotu datu modeli (5.2-5. attēls), kas novērš nepieciešamību konvertēt un tulcot projekta datu modeļus un grupēšanas noteikumus līdz kopējam saucējam. Tas ļauj grupēt datus pēc viena vai vairākiem atribūtiem, izmantojot tikai vienu koda rindu. Turpretī atvērtajā BIM un slēgtajā BIM, kur dati tiek glabāti daļēji strukturētos, parametriskos vai slēgtos formātos, apstrādei ir nepieciešamas desmitiemi vai pat simtiem kodu rindu un API izmantošana, lai mijiedarbotos ar ģeometrijas un atribūtu informāciju.

- ⌚ QTO strukturēta projekta grupēšanas piemērs pēc viena atribūta. Teksta vaicājums jebkurā LLM tērzēšanas telpā (ChatGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkurā citā):

Man ir CAD -projekts kā datu rāmis - lūdzu, filtrējet projekta datus, lai iegūtu vienumus ar parametru "Type", kas satur tikai "Type 1 vērtību. ↵

- LLM atbilde, visticamāk, būs Python koda veidā , izmantojot Pandas :

```
1 df[(df['Type'].str.contains("Type 1"))]
```

Run in IDE

**Attēls 5.2-6 Viena koda rindiņa, kas uzrakstīta ar LLM , ļauj sagrupēt visu CAD projektu pēc atribūta "Type" un iegūt vēlamo elementu grupu.**

Pateicoties 2D DataFrame vienkāršajai struktūrai, mums nav jāizskaidro LLM shēma un datu modelis, kas saīsina interpretācijas posmus un paātrina galīgo risinājumu izveidi. Agrāk pat vienkārša koda rakstīšanai bija jāapgūst programmēšanas valodas, bet tagad mūsdienīgi valodas modeļi (LLM) ļauj automātiski pārtulkot procesa loģiku kodā, strādājot ar strukturētiem datiem, izmantojot teksta vaicājumus.

LLM automatizācija un valodas modeļi var pilnībā novērst nepieciešamību speciālistiem, kas strādā ar CAD (BIM) datu grupēšanu un apstrādi, apgūt programmēšanas valodas vai BIM rīkus, nodrošinot iespēju risināt problēmas, izmantojot teksta pieprasījumus.

Tas pats pieprasījums - sagrupēt visus projekta elementus no kategorijas "sienas" un aprēķināt tilpumus katram tipam ( 5.2-5. attēls) -, kas CAD vidē (BIM) prasa 17 klikšķus saskarnē vai 40 rindu koda rakstīšanu, atvērtos datu apstrādes rīkos (piemēram, SQL vai Pandas) izskatās kā vienkāršs un intuitīvs pieprasījums:

- Ar vienu rindu Pandas:

```
df[df["Kategorija"].isin(["OST_Sienas"])]["Apjoms"].sum()
```

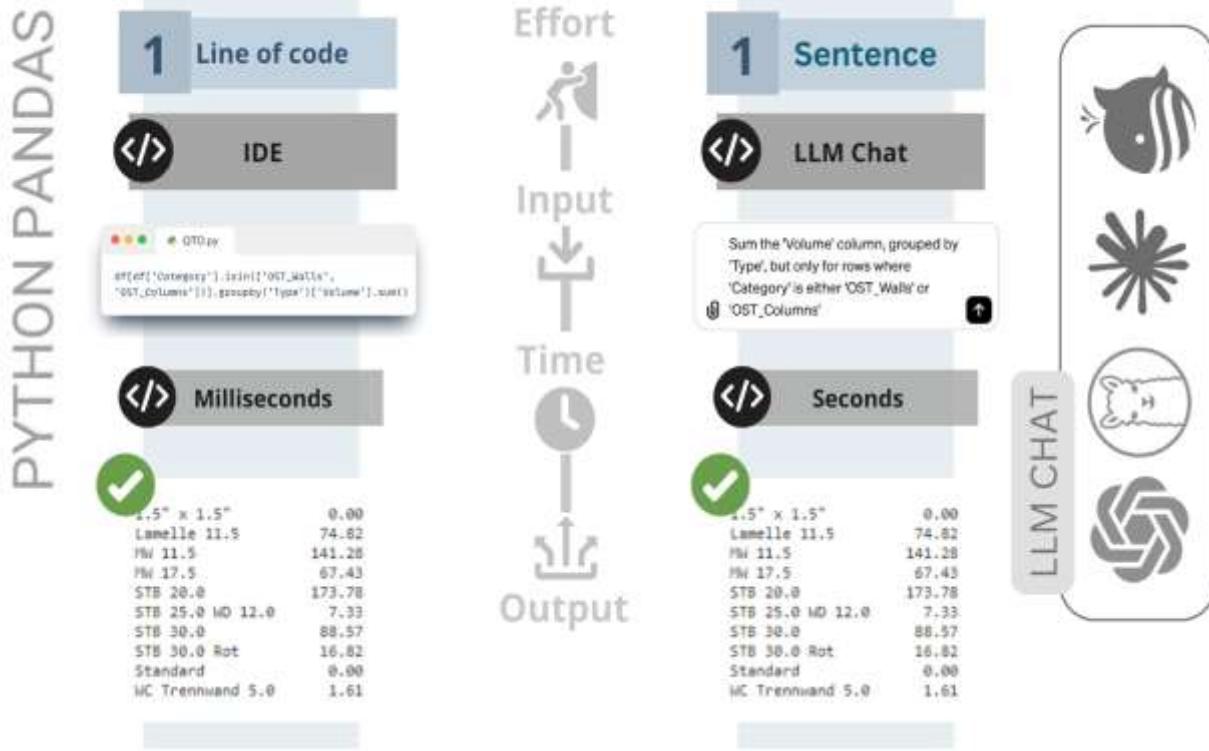
Kods dekodēšana: no df (DataFrame) ņem elementus, kuru atribūta slejā "Category" ir vērtības "OST\_Walls", sagrupē visus iegūtos elementus pēc atribūta slejas "Type" un iegūto elementu grupai summē atribūtu "Volume".

- Strukturēta projekta, kas iegūts no CAD, grupēšana, izmantojot SQL:

```
SELECT Veids, SUM(Apjoms) AS Kopējais apjoms
FROM elementi
WHERE Kategorija = 'OST_Walls'
GROUP BY Type;
```

- Ar LLM palīdzību mēs varam uzrakstīt grupēšanas pieprasījumu projekta datu bāzei kā vienkāršu teksta atsauci - uzvedni (5.2.2-7. attēls).:

Projekta datu rāmim sagrupējiet vienumus pēc parametra "Type", bet tikai vienumus, kuru parametrs "Category" ir vienāds ar "OST\_Walls" vai "OST\_Columns", un summējiet iegūtās kolonas parametru "Volume". ↵

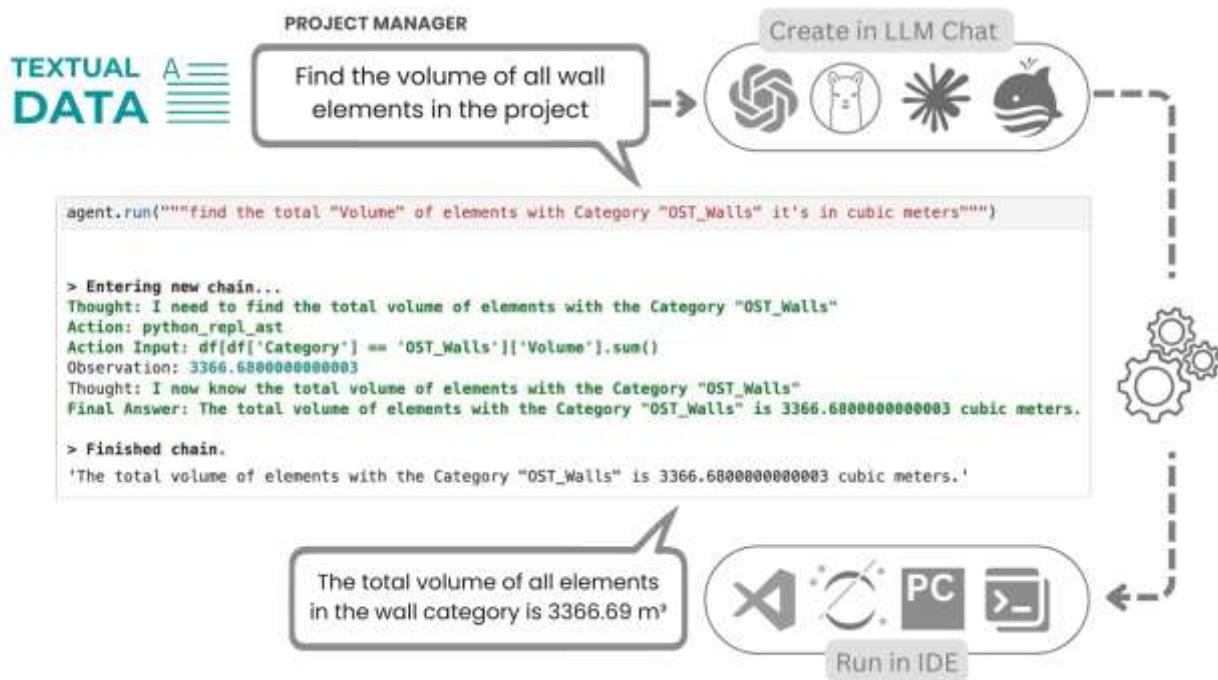


Attēls 5.2-7 Izmantojot SQL, Pandas un LLM, datu apstrādes automatizācija tagad ir iespējama, izmantojot dažas koda rindas un teksta vaicājumus.

QTO iegūšana no CAD datiem, izmantojot LLM rīkus (ChatGPT, LLaMa, Mistral, Claude, DeepSeek, Qwen, Grok), būtiski maina tradicionālās metodes atribūtu informācijas, kvantitatīvo un tilpuma datu iegūšanai atsevišķiem objektiem un objektu grupām.

Tagad pat projektu vadītāji, izmaksu aprēķinu vai logistikas speciālisti, kuriem nav padziļinātu zināšanu par projektēšanu un nav specializētas CAD programmatūras - (BIM-) pārdevēji, piekļūstot CAD datu bāzei, var dažu sekunžu laikā iegūt kopējo sienu vai citu objektu kategorijas elementu apjomu, vienkārši rakstot vai diktējot pieprasījumu.

Teksta vaicājumos ( 5.2-8. attēls) LLM aģents modelī apstrādā lietotāja pieprasījumu piemērot noteiktu funkciju vienam vai vairākiem parametriem - tabulas kolonnām. Rezultātā lietotājs sazinā ar LLM saņem vai nu jaunu kolonnu-parametru ar jaunām vērtībām, vai vienu konkrētu vērtību pēc grupēšanas



**Attēls 5.2-8 LLM modelis, kas strādā ar strukturētiem datiem, no teksta vaicājuma konteksta saprot, par kādu grupēšanu un atribūtiem lietotājs jautā.**

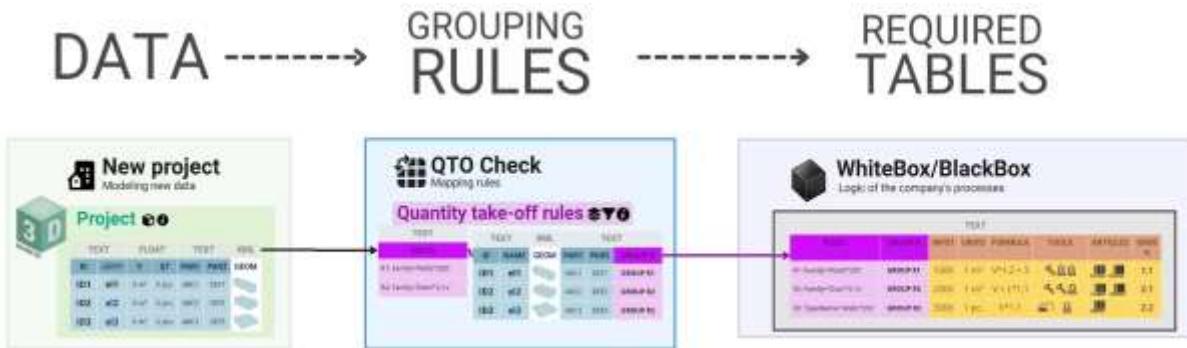
Ja ir nepieciešams iegūt daudzumus tikai vienai elementu grupai, ir pietiekami veikt vienkāršu QTO pieprasījumu ( 5.2-7. attēls) par CAD modeļa datiem. Tomēr, aprēķinot budžetu vai tāmi visam projektam, kas sastāv no daudzām elementu grupām, bieži ir nepieciešams iegūt daudzumus visiem elementu veidiem (klasēm0, kur katras elementu kategorija tiek apstrādāta atsevišķi - grupējot pēc attiecīgajiem atribūtiem).

Aplēsēju un vērtētāju praksē dažādiem objektu veidiem tiek izmantoti individuāli grupēšanas un aprēķinu noteikumi. Piemēram, logus parasti grupē pēc stāviem vai zonām (grupēšanas parametrs - atribūts Level, Rooms), bet sienas - pēc materiāla vai konstrukcijas tipa (parametrs Material, Type). Lai automatizētu grupēšanas procesu, šādi noteikumi ir iepriekš aprakstīti vietnē grupēšanas noteikumu tabulu veidā. Šīs tabulas darbojas kā konfigurācijas veidnes, kas nosaka, kādi atribūti jāizmanto

aprēķinos katrai elementu grupai projektā.

## **QTO visa projekta aprēķins, izmantojot grupas noteikumus no Excel izklājlapas.**

Reālos būvniecības projektos bieži vien ir nepieciešams veikt agregēšanu pēc vairākiem atribūtiem vienlaicīgi vienā elementu grupā. Piemēram, strādājot ar kategoriju "Windows" (kur atrībūtā Category ir tādas vērtības kā OST\_Windows vai IfcWindows), elementus var grupēt ne tikai pēc tipa - piemēram, pēc vērtības laukā Type Name vai Type -, bet arī pēc papildu raksturlielumiem, piemēram, pēc attiecīgajā atrībūtā norādītā siltumvadītspējas līmeņa. Šāda daudzdimensiju grupēšana ļauj iegūt precīzākus rezultātus konkrētai grupai. Līdzīgi, aprēķinot sienu vai grīdu kategorijas, kā filtri vai grupēšanas kritērijus var izmantot patvālīgas atrībūtu kombinācijas, piemēram, materiālu, līmeni, grīdu, ugunsizturību un citus parametrus (attēls 5.2-9).



Attēls 5.2-9 Katrai projekta vienību grupai vai kategorijai ir atšķirīga grupēšanas formula, kas sastāv no viena vai vairākiem kritērijiem.

Šādu grupēšanas noteikumu definēšanas process ir līdzīgs datu prasību izveides procesam, kas aprakstīts nodaļā "Prasību izveide un kvalitātes pārbaude dati" (4.4-5. attēls), kur tika detalizēti aplūkota darbs ar datu modeļiem. Šādi grupēšanas un aprēķinu noteikumi nodrošina rezultātu precīzitāti un atbilstību, lai automātiski aprēķinātu kopējos atrībūtus par vienību kategorijas daudzumu vai apjomu, nemot vērā visus nepieciešamos nosacījumus, kas jāņem vērā aprēķinos un aprēķinos.

- ➲ Tālāk dotois koda paraugs filtrē projektu tabulu tā, lai iegūtajā datu kopā būtu tikai tās vienības, kuru atrībūtu slejā "Category" ir vērtības "OST\_Windows" vai "IfcWindows" un vienlaikus atrībūtu slejā "Type" ir vērtība "Type 1":

Man ir projekts DataFrame - filtrēt datus tā, lai datu kopā palikuši tikai tie elementi, kuru atrībūtā "Category" ir vērtības "OST\_Windows" vai "IfcWindows" un tajā pašā laikā atrībūtā Type ir vērtība "Type 1" ↴

❓ LLM atbilde :



```
group.py
1 df[(df['Category'].isin(['OST_Windows', 'IfcWindows'])) & (df['Type'].str.contains("Type 1"))]
```

**Attēls 5.2-10 Viena koda rindiņa, kas ir līdzīga Excel formulai , ja už sagrupēt visas projekta vienības pēc vairākiem atribūtiem.**

legūtais kods (5.2-10. attēls) pēc CAD datu tulkošanas strukturētos atvērtos formātos ( 4. attēls).1-13) var palaist vienā no populārākajām IDE (integrētajām izstrādes vidēm), kuras minējām iepriekš, bezsaistes rezīmā: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populāros tiešsaistes rīkus: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

❓ Lai iegūtu projekta vienības QTO DataFrame veidlapā kategorijā "Windows" tikai ar konkrētu siltumvadītspējas vērtību, mēs varam izmantot šādu LLM vaicājumu:

Man ir projekts DataFrame - filtrēt datus tā, lai datu kopā paliktu tikai ieraksti ar "Category", kas satur "OST\_Windows" vai "IfcWindows" vērtības, un tajā pašā laikā slejā ThermalConductivity būtu vērtība 0.. ↴

❓ LLM atbilde :



```
group.py
1 df[(df['Category'].isin(['OST_Windows', 'IfcWindows'])) & (df['ThermalConductivity'] == 0.5)]
```

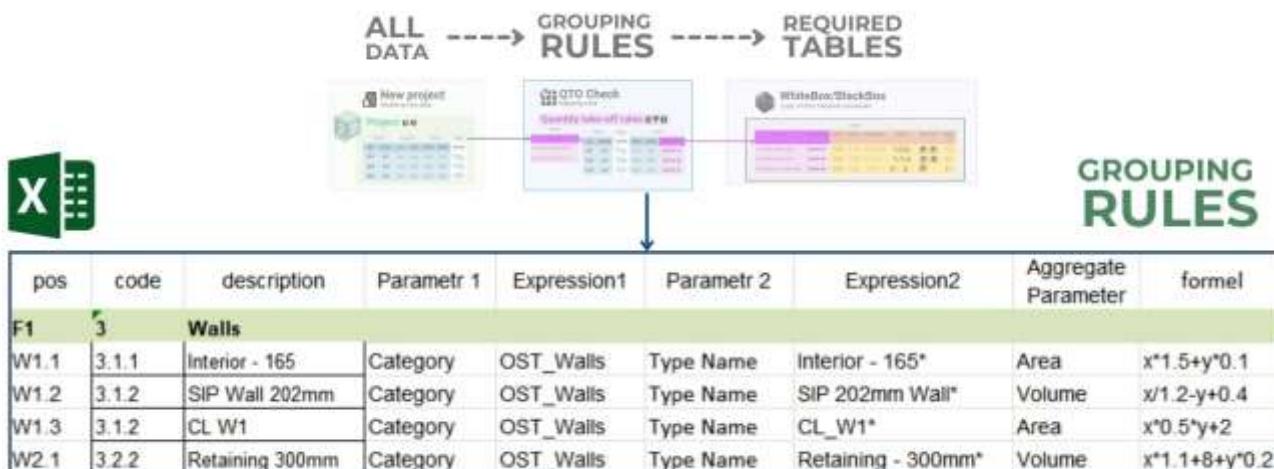
**Attēls 5.2-11 Ļoti vienkāršā Pandas vaicājumu valoda Python īauj vienlaicīgi palaist QTO jebkuram skaitam projektu.**

No LLM saņemtajā atbildē ( 5.2-11. attēls) loģiskais nosacījums "&" tiek izmantots, lai apvienotu divus kritērijus: siltumvadītspējas vērtību un piederību vienai no divām kategorijām. Ar metodi "isin" tiek pārbaudīts, vai atribūta slejas "Kategorija" vērtība ir norādītā sarakstā.

Projektos ar lielu elementu grupu skaitu un atšķirīgu grupēšanas loģiku - katrai projekta vienību kategorijai (piemēram, logiem, durvīm, plātnēm) jāizstrādā atsevišķi grupēšanas noteikumi, kas var ietvert papildu koeficientus vai kopējo atritūtu aprēķina formulas. Šajās formulās (attēls 5.2-12 atritūtu "formelēs", piemēram, x-daudzuma vērtība un y- grupas apjoms) un koeficientos nem vērā katras grupas unikālās īpašības, piem:

- % papildinājumi materiāla apjomam, lai ķemtu vērā pārsniegumus.
- fiksēts papildu materiāla daudzums
- korekcijas saistībā ar iespējamiem riskiem un aprēķinu klūdām formulās.

Kad filtrēšanas un grupēšanas noteikumi ir formulēti parametru formulu veidā katrai posteņu kategorijai, tos var saglabāt kā tabulu pa rindījām, piemēram, Excel formātā ( 5.2-12. attēls). Uzglabājot šos noteikumus strukturētā veidā, projekta datu ieguves, filtrēšanas un grupēšanas procesu var pilnībā automatizēt. Tā vietā, lai manuāli rakstītu daudzus atsevišķus vaicājumus, sistēma vienkārši nolasa parametru tabulu un piemēro attiecīgos noteikumus modelim (kopējam projekta datu rāmim ( 4.1-13. attēls)), ģenerējot galīgās QTO tabulas katrai projekta elementu kategorijai.



Attēls 5.2-12 QTO atritūtu grupēšanas tabula nosaka projekta elementu grupēšanas noteikumus, nodrošinot precīzu kopējo skaitu un apjomu katrai kategorijai.

Apkopotie noteikumi ļaus jums sagrupēt visu projektu un veikt visus nepieciešamos aprēķinus, tostarp apjoma atritūtu korekcijas. Rezultātā apjomi tiek pietuvināti "reālajam apjomam", kas tiek izmantots aprēķiniem un aprēķiniem, nevis tiem, kas sākotnēji atradās projektēšanas posmā CAD modelī.

Automātiski veidojot QTO apjoma tabulas visam projektam, lietojumprogrammai ir jāizskata visas grupēšanas noteikumu tabulas kategorijas, jāņem grupēšanas atribūti, pēc tiem jāgrupē visi projekta elementi un jāapkopo apjoma atribūts šai grupai, papildus reizinot to ar precīzējošo koeficientu vai koeficientu.

Palūgsim LLM uzrakstīt mums kodu šādam risinājumam, kur kodam būtu jāielādē divas tabulas - grupēšanas noteikumu tabula ( 5.2-12. attēls) un paša projekta datu tabula ( 4.1-13. attēls), pēc tam jāpiemēro grupēšanas noteikumi, jāsagrupē elementi atbilstoši dotajiem noteikumiem, jāaprēķina apkopotās vērtības un rezultāti jāsaglabā jaunā Excel failā.

❷ Nosūtiet teksta pieprasījumu uz LLM tērzēšanu:

Man ir nepieciešams kods, lai nolasītu projekta datus no faila 'basic\_sample\_project.xlsx' un pēc tam noteikumus no 'Grouping\_rules\_QTO.xlsx', ar kuriem sagrupēt visus datus pēc 'Parametrs 1' un 'Parametrs 2', apkopot 'Aggregate Parameter', filtrēt pēc 'Expression2', veikt aprēķinus no 'Formel1' un saglabāt QTO tabulu 'QTQ\_table2.xlsx'. ↵

■ LLM atbilde :



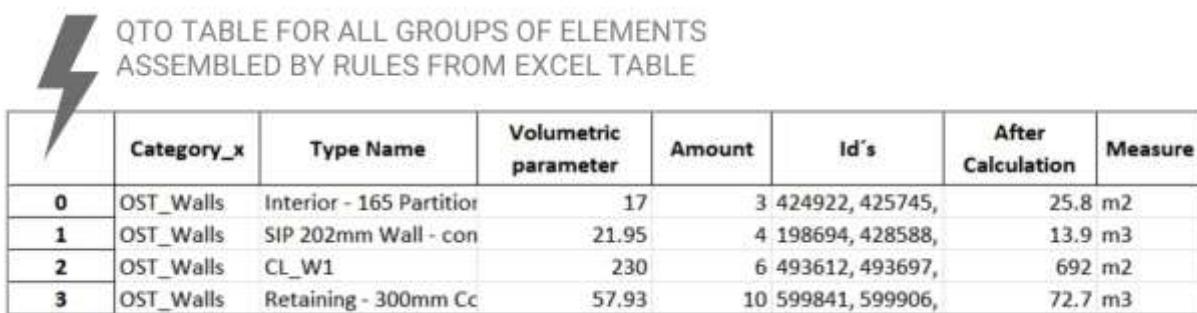
```

1 import pandas as pd
2
3 # Reading CSV and Excel files with project data and grouping rules respectively
4 df = pd.read_csv('../input/rac_basic_sample_project.xlsx', low_memory=False)
5 df_qto = pd.read_excel('../input//Grouping_rules_QTO.xlsx', header=0, skiprows=[1])
6
7 # Cleaning numerical columns in the dataframe by removing non-numeric characters
8 numeric_cols = ['Area', 'Volume', 'Depth', 'Width', 'Length', 'Perimeter']
9 df[numeric_cols] = df[numeric_cols]
10 df[numeric_cols] = df[numeric_cols].apply(lambda col: pd.to_numeric(col.astype(str).str.extract('(\d+\.\d+)')[0], errors='coerce'))
11
12 # Define a function to process each row in the grouping rules dataframe
13 def process_row(row):
14     # Group by specified parameters and calculate sum and count for the aggregate parameter
15     group = df.groupby([row["Parameter 1"], row["Parameter 2"]])
16     if row["Aggregate Parameter"] == "sum":
17         group = group[group.index.get_level_values(1).str.contains(row["Expression2"], regex=True)]
18
19     # Perform calculation defined in the rules
20     group['After Calculation'] = eval(row["formel"], {'x': group['sum'], 'y': group['count']})
21
22     return group.reset_index()
23
24 # Apply the function to each row in the grouping rules dataframe and concatenate the results
25 dfgroups = pd.concat([process_row(row) for _, row in df_qto.iterrows()]).reset_index(drop=True)
26
27 # Save the processed data to an Excel file using the OpenPyXL engine
28 dfgroups.to_excel('QTO_table2.xlsx', engine='openpyxl', index=False)

```

**Attēls 5.2-13 Python kods Pandas apstrādā projektēšanas datus no CAD projekta, sagrupējot tos saskaņā ar Excel programmā definētiem noteikumiem.**

Koda izpildes galarezultāts ( 5.2-13. attēls) būs vienību grupas tabula, kas satur ne tikai vispārinātos tilpuma atribūtus no sākotnējā CAD- (BIM-) modeļa, bet arī jaunu reālā tilpuma atribūtu, kurā nemitās vērā visas prasības pareizai aprēķinu un tāmju izveidei (piemērs 5.2-14. attēls).

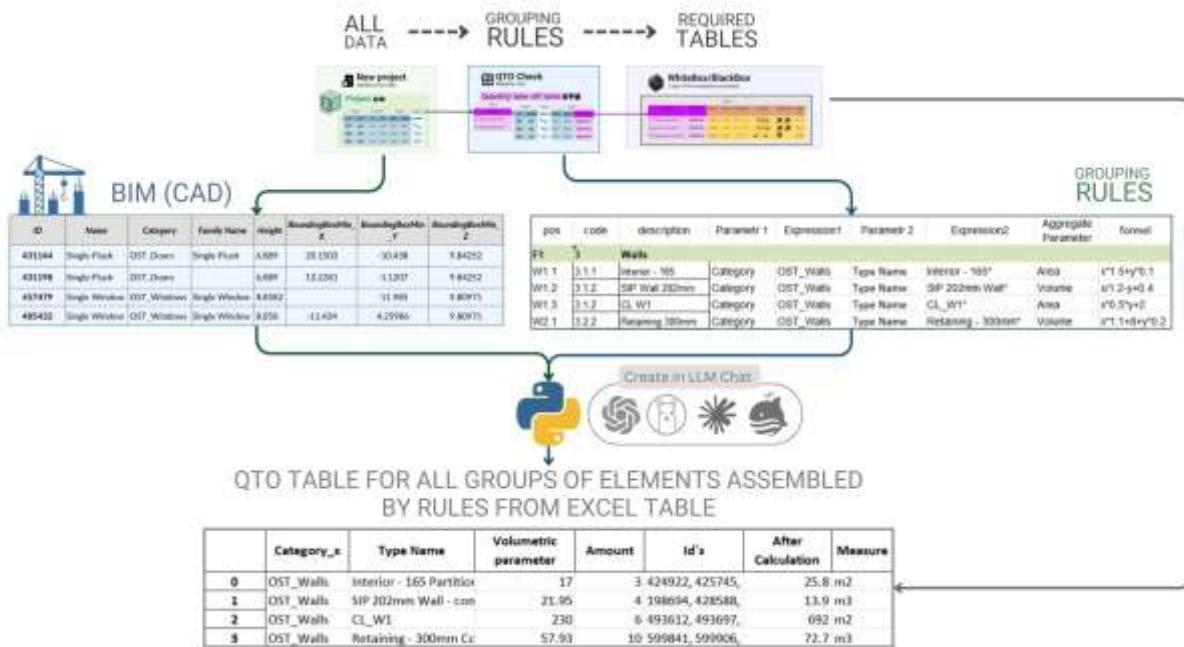


	Category_x	Type Name	Volumetric parameter	Amount	Id's	After Calculation	Measure
0	OST_Walls	Interior - 165 Partition		17	3 424922, 425745,		25.8 m2
1	OST_Walls	SIP 202mm Wall - con		21.95	4 198694, 428588,		13.9 m3
2	OST_Walls	CL_W1		230	6 493612, 493697,		692 m2
3	OST_Walls	Retaining - 300mm Cc		57.93	10 599841, 599906,		72.7 m3

**Attēls 5.2-14 Atribūts "Pēc aprēķina" tiek pievienots kopsavilkuma tabulai pēc tam, kad ir izpildīts kods, kas automātiski aprēķina faktisko apjomu.**

legūto kodu (5.2-13. attēls) var palaist kādā no populārākajām IDE (kuras minējām iepriekš) un piemērot jebkuram skaitam esošo vai jaunu ienākošo projektu (RVT, IFC, DWG, NWS, DGN u. c.), vai tie

būtu daži projekti vai varbūt simtiem projektu dažādos formātos strukturētā veidā (5.2-15. attēls).



Attēls 5.2-15 Automātiskais būvniecības datu grupēšanas process sasaista BIM datus (CAD) ar QTO tabulām, izmantojot Excel izklājlapas noteikumus.

Pielāgotais un parametrizētais tilpuma datu vākšanas process (5.2-15. attēls) ļauj pilnībā automatizēti vākt datus par projekta elementu kvantitatīvajiem raksturlielumiem un apjomiem, lai ar tiem varētu turpināt strādāt, tostarp aprēķināt izmaksas, logistiku, darba grafikus un aprēķināt oglekļa dioksīda pēdu, kā arī veikt citus analītiskus uzdevumus.

Apgūstot rīkus, kas ļauj viegli organizēt un grupēt projektu elementu grupas pēc noteiktiem atribūtiem, mēs tagad esam gatavi integrēt sagrupētus un filtrētus projektus ar dažādiem uzņēmuma aprēķiniem un biznesa scenārijiem.



## NODĀĻA 5.3.

### 4D, 6D -8D UN OGLEKĻA DIOKSĪDA EMISIJU APRĒĶINS CO<sub>2</sub>.

#### 4D modelis: laika integrēšana būvniecības aplēsēs

Papildus izmaksu aprēķināšanai viens no galvenajiem projektēšanas datu lietojumiem būvniecībā ir laika parametru noteikšana - gan atsevišķām būvniecības operācijām, gan visam projektam kopumā. Uz resursiem balstītu aplēšu metodi un ar to saistīto aprēķinu datu bāzi, kas tika detalizēti aplūkota iepriekšējā nodaļā "Aprēķini un tāmes būvprojektiem", bieži izmanto kā pamatu automatizētam laika aprēķinam un darbu izpildes grafika izveidei.

Uz resursiem balstītā pieeja ļem vērā ne tikai materiālu izmaksas, bet arī laika resursus. Aprēķinot izmaksas, katram procesam var piešķirt darba uzdevuma atribūtu ( 5.3-1. attēls - Darba uzdevuma parametrs), kā arī laika daudzumu un izmaksas, kas saistītas ar šī procesa izpildi. Šie parametri ir īpaši svarīgi, lai aprakstītu darbības, kurām nav noteiktas tirgus cenas un kuras nav tieši nopērkamas, piemēram, būviekkārtu izmantošana, strādnieku nodarbināšana vai loģistikas procesi (kurus parasti parasti izsaka stundās). Šādos gadījumos izmaksas nenosaka iepirkuma departaments, bet gan tieši īstenošanas uzņēmums, pamatojoties uz iekšējām normām vai ražošanas likmēm ( 5.3-1. attēls).



Attēls 5.3-1. Darbu aprēķini, izmantojot uz resursiem balstītu novērtēšanas metodi, ietver darba stundu laika grafiku.

Tādējādi aprēķinos izmaksu aprēķina līmenī ir iekļautas ne tikai degvielas un materiālu izmaksas (iegādes izmaksas), bet arī autovadītāju, tehniku un palīgstrādnieku pavadītais laiks uz vietas. Dotajā

piemērā ( 5.3-1. att.) izmaksu tabulā ir aprēķinātas izmaksas par pamatu bloka uzstādīšanu, ietverot tādus darbu veidojošos posmus kā sagatavošana, karkasa uzstādīšana un betonēšana, kā arī nepieciešamos materiālus un darbaspēka izmaksas. Tomēr atsevišķām darbībām, piemēram, sagatavošanas darbiem, var nebūt materiālu izmaksu, bet var būt ievērojamas pagaidu darbaspēka izmaksas, kas izteiktas cilvēkstundās.

Lai plānotu darbu secību (darba grafikam) būvlaukumā, aprēķinu tabulai manuāli pievieno atribūtu "Darba pasūtījums" ( 5.3-1. attēls). Šo atribūtu norāda papildu ailē tikai tiem posteņiem, kuru mērvienība ir izteikta laikā (stundā, dienā). Šis atribūts papildina darba kodu, aprakstu, daudzumu, mērvienību (parametrs "Vienība") un izmaksas. Darbību skaitiskā secība (parametrs "Darbu secība") ļauj noteikt secību, kādā būvlaukumā tiek veikti uzdevumi, un izmantot to plānošanai.

## **Būvdarbu grafiks un tā automatizācija, pamatojoties uz izmaksu aprēķina datiem**

Būvdarbu grafiks ir vizuāls priekšstats par darbu plānu un procesiem, kas jāveic projekta īstenošanas gaitā. Tas tiek veidots, pamatojoties uz detalizētiem resursu aprēķiniem ( 5.3-1. attēls), kur katrs uzdevums-darbs tiek plānots, papildus resursu izmaksām, pēc laika un secības.

Atšķirībā no vidējās vērtības noteikšanas metodēm, kur laika aplēses balstās uz materiālu vai iekārtu uzstādīšanas tipiskām stundām, uz resursiem balstītajā metodē plānošana balstās uz faktiskajiem datiem izmaksu aprēķinā. Katrs ar darbaspēka izmaksām saistītais tāmes postenis ir balstīts uz piemēroto kalendāru, kurā ķemti vērā faktiskie resursu izmantošanas nosacījumi darba periodā. Ražošanas stundu koriģēšana, izmantojot koeficientus izmaksu aprēķina līmenī (attēls 5.3-1 parametrs "Bid. koeficients"), ļauj ķemt vērā produktivitātes atšķirības un sezonālās īpatnības, kas ietekmē darba grafiku.

Lai noteiktu procesa sākuma un beigu datumus būvdarbu grafikam Ganta diagrammā, mēs ķemam katras elementa laika apjoma atribūta vērtības no pamatu bloku izmaksu aprēķina un reizinām tās ar bloku skaitu (šajā gadījumā ar betona pamatu bloku skaitu). Šis aprēķins dod katras uzdevuma ilgumu. Pēc tam mēs uzzīmējam šos ilgumus uz laika līnijas, sākot no projekta sākuma datuma, lai izveidotu grafiku, un rezultāts ir vizuāls attēlojums, kas parāda, kad katram uzdevumam ir jāsākas un jābeidzas. Procesu parametrs "Darba kārtība" papildus ļauj mums saprast, vai darba process notiek paralēli ("Darba kārtība", piemēram, 1.1-1.1) vai secīgi (1.1-1.2).

**Ganta diagramma** ir grafisks rīks plānošanai un projektu vadībai, kas attēlo uzdevumus kā horizontālas joslas uz laika grafika. Katras josla norāda uzdevuma ilgumu, tā sākumu un beigas.

Grafiks jeb Ganta diagramma palīdz projekta vadītājiem un strādniekiem skaidri saprast, kad un kādā secībā jāveic dažādi būvniecības posmi, nodrošinot efektīvu resursu izmantošanu un termiņu ievērošanu.

Sniegsim grafiku trīs betona pamatu bloku montāžai, izmantojot aprēķinus no iepriekš dotās tabulas. Izmantojot izmaksu tabulu ( 5.3-1. attēls) no iepriekš dotā piemēra, lūdziet LLM sastādīt grafiku 3

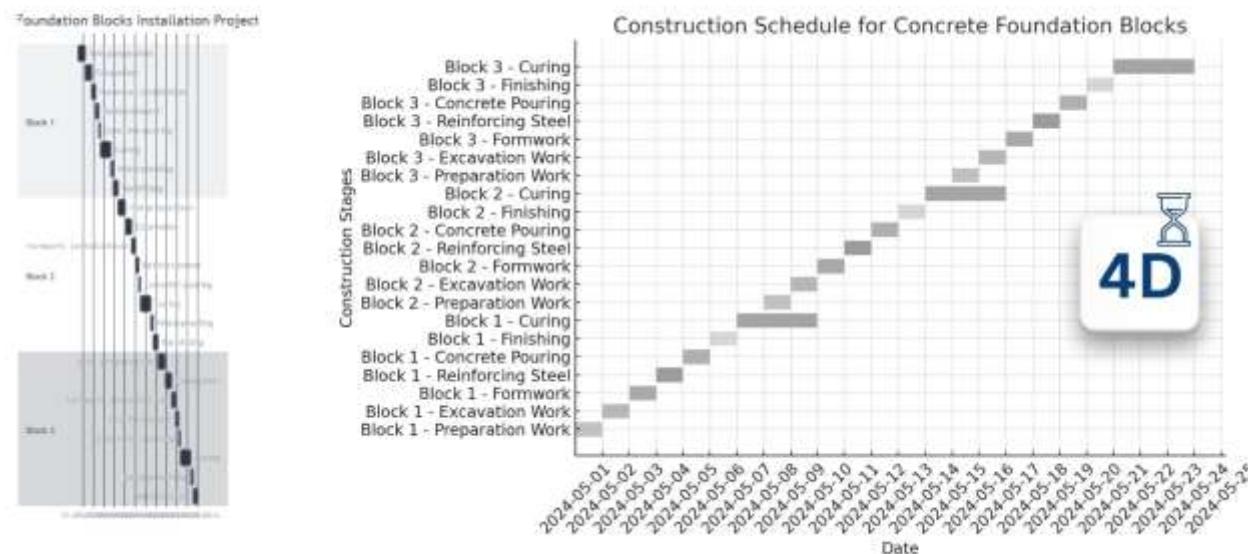
pamatu bloku elementu uzstādīšanai, piemēram, 2024. gada 1. maijā.

Lai nosūtītu izmaksu aprēķinu uz LLM, mēs varam augšupielādēt izmaksu aprēķina tabulu XLSX formātā vai vienkārši ievietot izmaksu aprēķina JPEG attēla ekrānšāviņu tieši LLM tērzēšanā ( 5.3-1. attēls). LLM patstāvīgi atradīs bibliotēku tabulas attēla vizualizēšanai un, reizinot darbību laika atribūtus no tabulas ar to apjomu, saliks visus datus grafikā.

❓ Nosūtiet teksta pieprasījumu uz LLM:

Izveidojiet darba grafiku Ganta diagrammas veidā, izmantojot ar laiku saistītās tabulas vērtības ( 5.3-1. attēls kā JPEG pielikumā) 3 pamatu bloku uzstādīšanai. Darbi katram blokam jāveic secīgi. Darbu sākumu nosaka no 01.05.2024. gada. ↵

❓ LLM atbilde



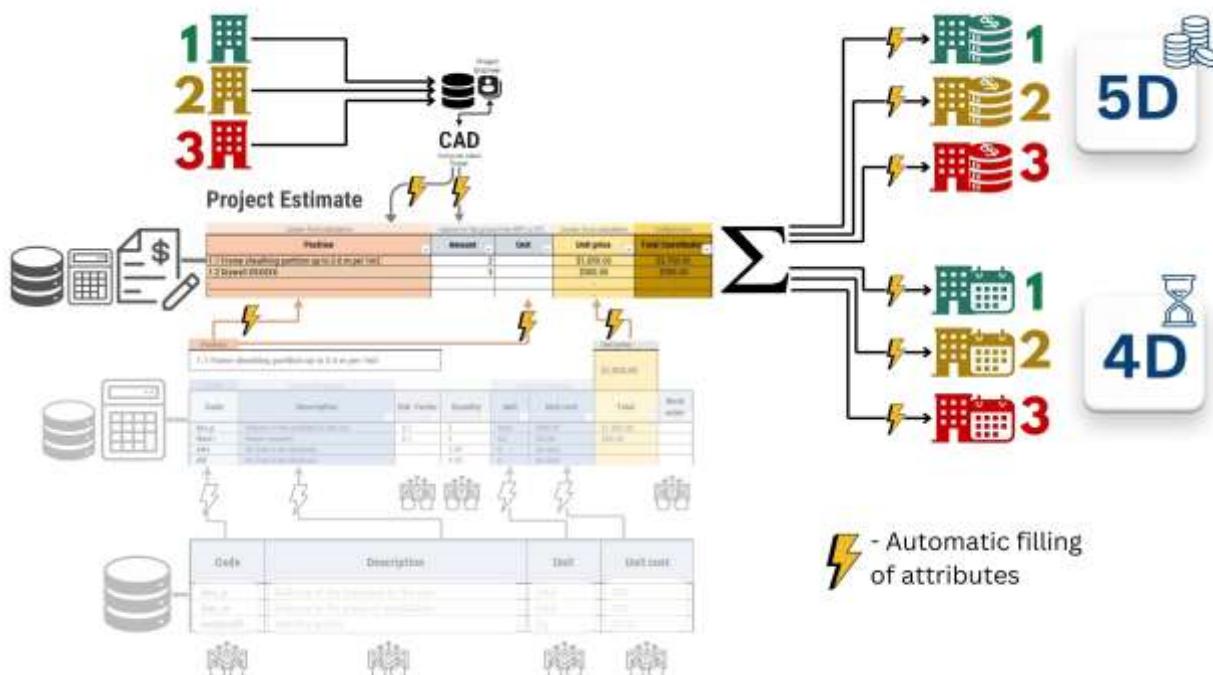
**Attēls 5.3-2 Automātiski ģenerē vairāki LLM Ganta diagrammā ir parādīti trīs betona bloku būvniecības posmi saskaņā ar izaicinājumā dotajiem nosacījumiem.**

legūtais grafiks ( 5.3-2. attēls) ir laika diagramma, kurā katrs horizontālais stabiņš atbilst noteiktam darbu posmam pie pamatu bloka un parāda operāciju secību (parametrs "Darbu secība"), piemēram, sagatavošana, rakšana, veidņu montāža, pastiprināšana, betona ieliešana un apdare, t.i., tie procesi, kuriem aprēķinos ir aizpildīti laika parametri un secība.

Šādā grafikā ( 5.3-2. attēls) nav ņemti vērā ierobežojumi, kas saistīti ar darba dienām, maiņām vai darba laika normām, bet tas ir paredzēts tikai konceptuālai procesa vizualizācijai. Precīzu grafiku, kas atspoguļos darba paralēlismu, var papildināt ar atbilstošiem pamācībām vai papildu norādījumiem tērzēšanas telpā

Izmantojot vienotu izmaksu aprēķinu ( 5.3-1. attēls), pateicoties tilpuma atribūtiem no 3D -ģeometrijas, ir iespējams automātiski aprēķināt gan projekta izmaksas, izmantojot automatizētus aprēķinus, gan

vienlaikus aprēķināt grupu laika raksturlielumus tabulu vai grafiku veidā dažādiem projekta variantiem ( 5.3-3. attēls).



Attēls 5.3-3 Automātiskais aprēķins, ļauj veikt tūlīteju un automātisku izmaksu un laika prognozi dažādiem projekta variantiem.

Modernās modulārās ERP -sistēmas ( 5.4-4. attēls), ielādējot datus no CAD modeļiem, izmanto līdzīgas automatizētas laika aprēķina metodes, kas ievērojami samazina lēmumu pieņemšanas procesu. Tas ļauj nekavējoties un precīzi plānot darbu grafikus un aprēķināt kopējo laiku, kas nepieciešams visu projekta realizācijas uzdevumu izpildei, nemot vērā reālās cenas.

### Paplašinātie atribūtu slāņi 6D -8D: no energoefektivitātes līdz drošības nodrošināšanai

6D, 7D un 8D ir paplašināti informācijas modelēšanas līmeņi, no kuriem katrs sniedz papildu atribūtu slāņus visaptverošajam projekta informācijas modelim, kura pamatā ir 3D -modeļa atribūti ar to skaitu un apjomu. Katrs papildu līmenis ievieš ūpašus parametrus, kas nepieciešami turpmākai grupēšanai vai turpmākai identifikācijai citās sistēmās, piemēram, ūpašuma pārvaldības sistēmās (PMS), datorizētajā telpu pārvaldībā (CAFM), būvniecības projektu pārvaldībā (CPM) un drošības pārvaldības sistēmās (SMS).



**Attēls 5.3-4. Attribūti 6D, 7D un 8D datu informācijas modelī paplašina dažādu projekta aspektu  
- no energoefektivitātes līdz drošībai - izskatīšanu.**

- **6D** papildus projekta datu bāzei (vai datu rāmim ( 4.1-13. attēls)) ar elementu ģeometriskajiem un tilpuma atribūtiem tiek pievienota informācija (atribūtu kolonas) par vides ilgtspējību. Tas ietver informāciju, kas saistīta ar energoefektivitāti, oglekļa dioksīda emisiju, materiālu pārstrādājamību un videi draudzīgu tehnoloģiju izmantošanu. Šie dati jauj novērtēt projekta ietekmi uz vidi, optimizēt projekta lēmumus un sasniegt ilgtspējīgas attīstības mērķus (ESG).
- **7D** atribūti papildina atribūtus, kas nepieciešami ēku uzturēšanas pārvaldībai. Tie ir dati par tehniskās apkopes grafikiem, sastāvdaļu dzīves cikliem, tehnisko dokumentāciju un remonta vēsturi. Šis informācijas kopums nodrošina, ka modeli var integrēt ar tehniskās apkopes sistēmām (CAFM, AMS), jauj efektīvi plānot tehnisko apkopi, iekārtu nomaiņu un nodrošina atbalstu visā ēkas dzīves ciklā.
- **8D** papildu atribūtu slānis, - ietver ar drošību saistītu informāciju - gan būvniecības posmā, gan turpmākajā ekspluatācijā. Šajā modelī iekļauti personāla drošības pasākumi, instrukcijas ārkārtas situācijām, evakuācijas sistēmas un ugunsdrošības prasības. Šo datu integrācija digitālajā modelī palīdz iepriekš ķemt vērā riskus un izstrādāt arhitektūras, inženiertehniskos un organizatoriskos risinājumus, kuros ķemtas vērā veselības un drošības prasības.

Strukturētā tabulas formā 4D līdz 8D slāņos ir attēloti papildu atribūti kolonu veidā ar aizpildītām vērtībām ( 5.3-5. attēls), kas pievienoti jau aizpildītajiem 3D - modeļa atribūtiem, piemēram, nosaukumam, kategorijai, tipam un tilpuma raksturojumam. Vērtības atribūtu slānos 6D, 7D un 8D satur papildu teksta un skaitliskos datus, piemēram, pārstrādes procentuālo daļu, oglekļa dioksīda emisijas nospiedumu, garantijas periodu, nomaiņas ciklu, uzstādišanas datumu, drošības protokolus utt.

ID	Type Name	Width	Length	Recyclability	Carbon Footprint	Warranty Period	Replacement Cycle	Maintenance Schedule	Installation Date	Wellbeing Factors	Safety Protocols
W-NEW	Window	120 cm	-	90%	1622 kgCO <sub>2</sub> e	8 years	20 years	Annual	-mon	XYZ Windows	ISO 45001
W-OLD1	Window	100 cm	140 cm	90%	1522 kgCO <sub>2</sub> e	8 years	15 years	Biannual	08/22/2024	XYZ Windows	OSHA Standard
W-OLD2	Window	110 cm	160 cm	90%	1522 kgCO <sub>2</sub> e	-	15 years	Biannual	08/24/2024	????	OSHA Standard
D-122	Door	90 cm	210 cm	100%	1322 kgCO <sub>2</sub> e	15 years	25 years	Biennial	08/25/2024	Doors Ltd.	OSHA Standard

Attēls 5.3-5 6D -8D pievieno atribūtu slāņus datu informācijas modelim, kas jau satur ģeometriskos un tilpuma atribūtus no 3D -modeļa.

Mūsu jaunajam logam ( 4.4-1. attēls) elementam ar identifikatoru W-NEW ( 5.3-5. attēls) var būt šādi 3D -8D atribūti:

**3D** -atribūti - ģeometriskā informācija, kas iegūta no CAD sistēmām:

- "Tipa nosaukums" - elements "Window"
- "Platums" - 120 cm
- Turklat kā atsevišķu atribūtu var pievienot elementa "Bounding Box " punktus vai tā "ģeometrijas BREP / MESH ".

**6D** atribūti - vides ilgtspēja:

- 90 procentu pārstrādes līmenis
- "Oglekļa pēda" - 1 622 kg CO<sub>2</sub>

**Atribūti 7D** - objektu pārvaldības dati:

- "Garantijas termiņš" - 8 gadi
- "Rezerves cikls" ir 20 gadus vecs
- "Uzturēšana" - nepieciešama katru gadu

**8D** atribūti - ēku drošas lietošanas un ekspluatācijas nodrošināšana:

- "Instalēts" logs - uzņēmums "XYZ Windows"
- "Drošības standarts" - atbilst standartam ISO 45001

Visi datu bāzē vai datu kopā reģistrētie parametri ( 5.3-5. attēls) ir nepieciešami dažādu nodaļu speciālistiem grupēšanai, meklēšanai vai aprēķiniem. Šāds daudzdimensiju uz atribūtiem balstīts projekta objektu apraksts sniedz pilnīgu priekšstatu par to dzīves ciklu, ekspluatācijas prasībām un daudziem citiem aspektiem, kas nepieciešami projekta projektēšanai, būvniecībai un ekspluatācijai.

## CO<sub>2</sub> aplēses un oglekļa dioksīda emisiju aprēķins būvniecības projektos

Līdztekus tēmai par būvniecības projektu ilgtspēju 6D posmā (attēls 5.3-5), mūsdienu būvniecībā galvenā uzmanība tiek pievērsta projektu vides ilgtspējai, kur par vienu no galvenajiem aspektiem kļūst oglekļa dioksīda CO<sub>2</sub> emisiju, kas rodas projekta dzīves cikla posmos (piemēram, ražošanā un uzstādīšanā), novērtēšana un samazināšana.

Būvmateriālu oglekļa emisiju aplēse un aprēķins ir process, kurā kopējās oglekļa emisijas nosaka, reizinot projektā izmantotā elementa vai elementu grupas tilpuma rādītājus ar attiecīgajai kategorijai piemērotu oglekļa emisiju faktoru.

Oglekļa emisiju ņemšana vērā, vērtējot būvniecības projektus kā daļu no plašākiem ESG kritērijiem (vides, sociālie un pārvaldības kritēriji), analīzei piešķir jaunu sarežģītības pakāpi. Tas ir īpaši svarīgi klientam-investētājam, lai iegūtu attiecīgu sertifikātu, piemēram, LEED® (Leadership in Energy and Environmental Design), BREEAM® (Building Research Establishment Environmental Assessment Method) vai DGNB® (Deutsche Gesellschaft für Nachhaltiges Bauen). Iegūstot kādu no šiem sertifikātiem, var ievērojami palielināt īpašuma pārdošanas iespējas, vienkāršot nodošanu ekspluatācijā un nodrošināt atbilstību uz ilgtspēju orientētu īrnieku (ESG) prasībām. Atkarībā no projekta prasībām var izmantot arī HQE (Haute Qualité Environnementale, Francijas zaļo ēku standarts), WELL (WELL Building Standard, kas vērts uz lietotāju veselību un komfortu) un GRESB (Global Real Estate Sustainability Benchmark)

Vides, sociālā un pārvaldības **ESG** (environmental, social and governance) ir plašs principu kopums, ko var izmantot, lai novērtētu uzņēmuma pārvaldību, sociālo un vides ietekmi gan iekšēji, gan ārēji.

ESG, ko sākotnēji 2000. gadu sākumā izstrādāja finanšu fondi, lai sniegtu ieguldītājiem informāciju par plašiem vides, sociālajiem un pārvaldības kritērijiem, ir kļuvis par galveno rādītāju gan uzņēmumu, gan projektu, tostarp būvniecības projektu, novērtēšanā. Saskaņā ar lielāko konsultāciju uzņēmumu veikto pētījumu datiem vides, sociālie un pārvaldības (ESG) apsvērumi kļūst par neatņemamu būvniecības nozares sastāvdaļu.

Saskaņā ar EY (2023) "Ceļš uz oglekļa neutralitāti" (The Path to Carbon Neutrality), uzņēmumi, kas aktīvi īsteno ESG principus, ne tikai samazina ilgtermiņa riskus, bet arī palielina savu uzņēmējdarbības modeļu efektivitāti, kas ir īpaši svarīgi globālās tirgus transformācijas apstākļos [103]. PwC ziņojumā par ESG informētību norādīts, ka uzņēmumu informētība par ESG faktoru nozīmi svārstās no 67% līdz 97%, un lielākā daļa organizāciju uzskata, ka šīs tendences ir būtiskas nākotnes ilgtspējas nodrošināšanai [104], un ka uzņēmumi lielākoties novēro ievērojamu ieinteresēto personu spiedienu integrēt ESG principus.

Tādējādi ESG -principu integrēšana būvniecības projektos ne tikai palīdz iegūt starptautiskus ilgtspējas sertifikātus, piemēram, LEED, BREEAM, DGNB, bet arī nodrošina nozares uzņēmumu ilgtermiņa ilgtspēju un konkurētspēju.

Viens no būtiskākajiem faktoriem, kas ietekmē būvniecības projekta kopējo oglekļa dioksīda emisiju, ir būvniecības materiālu un komponentu ražošanas un loģistikas posmi. Būvlaukumā izmantotajiem materiāliem bieži vien ir izšķiroša ietekme uz kopējo CO<sub>2</sub> emisiju apjomu, jo īpaši projekta dzīves cikla sākumposmā - no izejvielu ieguves līdz piegādei uz būvlaukumu.

Lai aprēķinātu emisijas pa ēku elementu kategorijām vai veidiem, ir jāizmanto oglekļa emisijas standartkoeficienti, kas atspoguļo CO<sub>2</sub> daudzumu, kas rodas, ražojot dažādus materiālus. Šādi materiāli ir betons, kieģeļi, pārstrādāts tērauds, alumīnijs un citi. Šīs vērtības parasti tiek iegūtas no atzītiem avotiem un starptautiskām datubāzēm, piemēram, UK ICE 2015 (Inventory of Carbon and Energy) un US EPA 2006 (U.S. Environmental Protection Agency) [105]. Turpmākajā tabulā (5.3-6. attēls) apkopoti bāzes emisiju faktori virknei izplatītu būvmateriālu. Katram materiālam ir norādīti divi galvenie parametri: īpatnējās CO<sub>2</sub> emisijas (kilogramos uz kilogramu materiāla) un pārrēķina koeficienti no tilpuma uz svaru (kilogramos uz kubikmetru), kas nepieciešami, lai integrētu aprēķinus projektēšanas modelī un sasaistītu ar QTO datu grupēšanu.



Carbon Emitted in Production		UK ICE Database (2015) USEPA (2006)	UK ICE Database (2015) USEPA (2006)	Coefficient m <sup>3</sup> to kg
Material	Abbreviated	Process Emissions (kg CO <sub>2</sub> e/ kg of product) (K1)	Process Emissions (kg CO <sub>2</sub> e/ kg of product) (K2)	Kg / m <sup>3</sup> (K3)
Concrete	Concrete	0.12	0.12	2400
Concrete block	Concrete_block	0.13**	0.14	2000
Brick	Brick	0.24	0.32	2000
Medium density fiberboard (MDF)	MDF	0.39*	0.32	700
Recycled steel (avg recy content)	Recycled_steel	0.47	0.81	7850
Glass (not including primary mfg.)	Glass	0.59	0.6	2500
Cement (Portland, masonry)	Cement	0.95	0.97	1440
Aluminum (virgin)	Aluminum	12.7%	16,6	2700

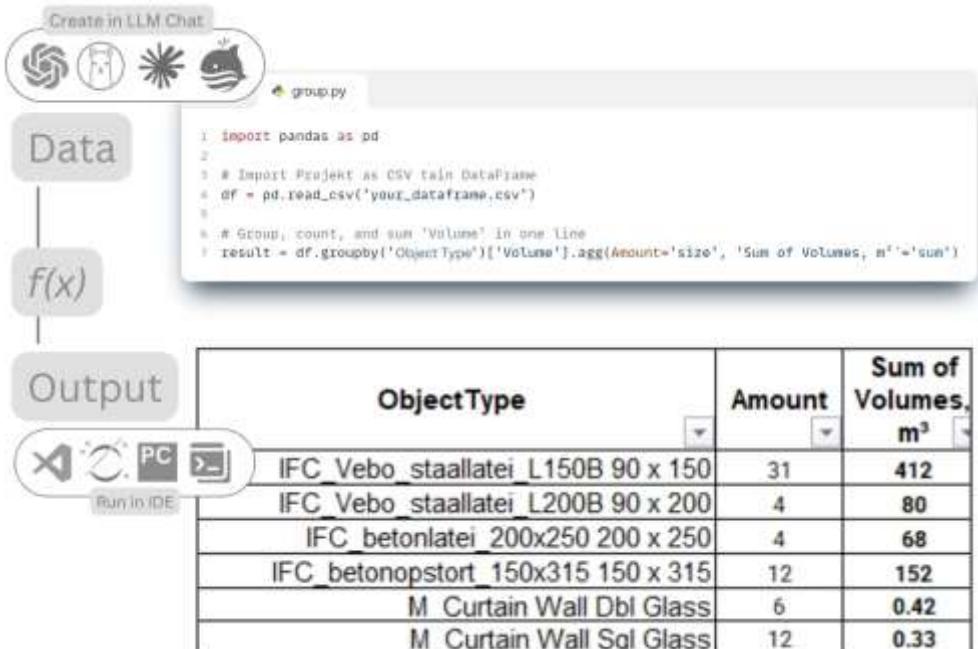
Attēls 5.3-6 Dažādu būvmateriālu ražošanas laikā emitētā oglekļa daudzums saskaņā ar Apvienotās Karalistes ICE un ASV EPA datubāzes datiem.

Lai aprēķinātu projekta kopējās CO<sub>2</sub> emisijas, tāpat kā 4D un 5D aprēķinos, ir jānosaka katras objektu grupas atribūtu apjomī. To var izdarīt, izmantojot kvantitatīvās analīzes rīkus (QTO), iegūstot atribūtu apjomus kubikmetros, kā sīkāk aplūkots sadaļā par daudzuma ņemšanu. Pēc tam šos apjomus reizina ar attiecīgajiem koeficientiem katras materiālu grupas atribūtam "CO<sub>2</sub> procesa emisijas".

- Automātiski iegūsim tilpumu tabulu pēc elementu tipa no CAD (BIM) projekta, sagrupējot visus projekta datus, kā tas jau tika darīts iepriekšējās nodalās. Lai veiktu šo uzdevumu, izmantosim LLM.

Lūdzu, sagrupējiet datu rāmja tabulu no CAD (BIM) projekta pēc kolonnas parametra "Objekta nosaukums" (vai "Tips") un norādiet vienību skaitu katrā grupā, kā arī apkopojiet parametru "Apjoms" visiem attiecīgā tipa vienībām. ↴

❓ LLM atbilde :

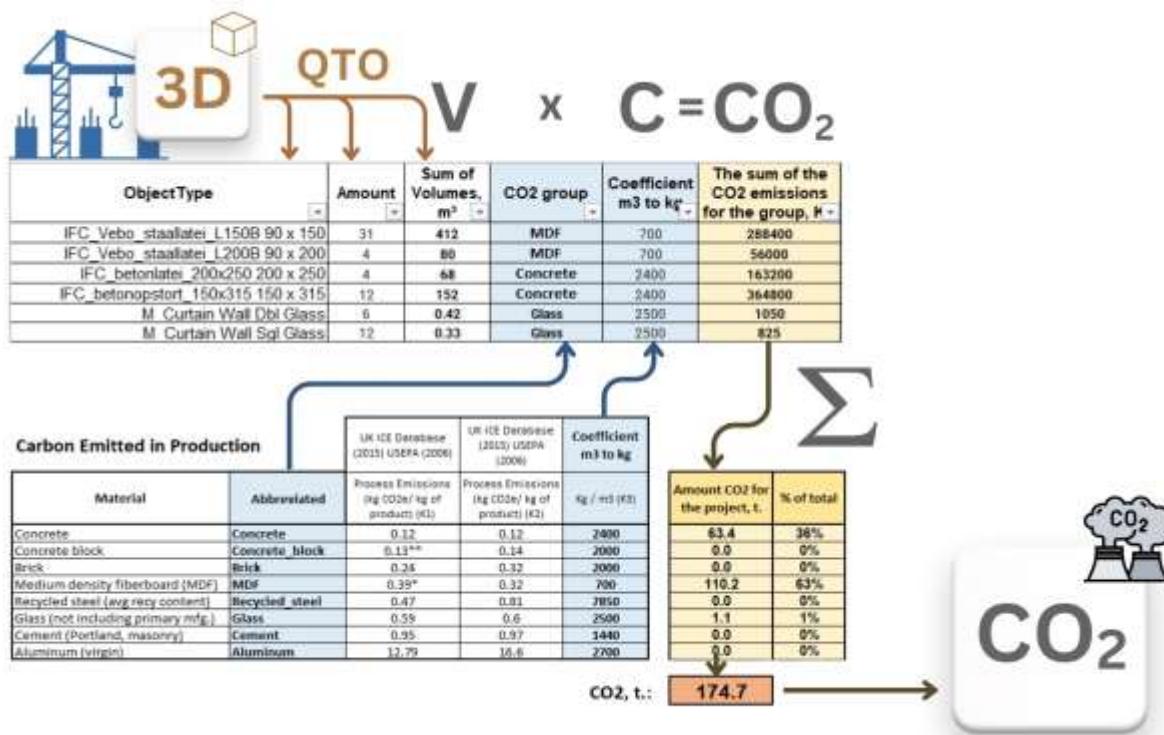


**Attēls 5.3-7 LLM ģenerētais kods mums ir sagrupējis projekta vienības pēc tipa (ObjectType) ar kopsavilkuma atribūtu "Volume".**

Lai automatizētu kopējo CO<sub>2</sub> emisiju aprēķinu visam projektam, vienkārši iestatiet automātisku datu kartēšanu tabulā vai manuāli sasaistiet elementu tipus ( 5.3-7. attēls) ar atbilstošajiem materiālu tipiem ( 5.3-6. attēls) no emisijas faktoru tabulas. Gatavu tabulu ar emisijas koeficientiem un formulām, kā arī kodu, lai iegūtu apjomus no CAD formātiem (BIM) un automatizētu CO<sub>2</sub> noteikšanu, var atrast GitHub, meklējot "CO<sub>2</sub>\_calculating-the-embodyied-carbon. DataDrivenConstruction." [106].

Tādējādi datu integrācija pēc QTO elementu grupēšanas no CAD datubāzes jauj automātiski aprēķināt oglēkļa dioksīda emisijas ( 5.3-8. attēls) dažādiem konstrukcijas variantiem. Tas jauj analizēt dažādu materiālu ietekmi dažādos variantos un izvēlēties tikai tos risinājumus, kas atbilst CO<sub>2</sub> emisiju prasībām, lai nododot ēku ekspluatācijā, iegūtu konkrētu sertifikātu

CO<sub>2</sub> emisiju aplēse, reizinot koeficientus ar sagrupēto projekta elementu apjomiem, ir tipisks piemērs uzdevumam, ko veic būvniecības uzņēmums, lai objektam iegūtu ESG novērtējumu (piemēram, LEED sertifikātu).



Attēls 5.3-8 QTO grupu integrācija no CAD datubāzēm nodrošina precizitāti un automatizāciju galīgo CO<sub>2</sub> emisiju aprēķinu iegūšanā.

Līdzīgi, definējot elementu grupu apjomus, mēs varam veikt aprēķinus materiālu kontrolei un loģistikai, kvalitātes uzraudzībai un pārvaldībai, enerģijas modelēšanai un analīzei un daudziem citiem uzdevumiem, lai iegūtu jaunu atribūta statusu (parametru tabulā) gan atsevišķām elementu grupām, gan visam projektam.

Ja šādu aprēķinu procesu skaits uzņēmumā sāk pieaugt, rodas jautājums par nepieciešamību automatizēt šādus aprēķinus un ieviest aprēķinu rezultātus uzņēmuma procesos un datu pārvaldības sistēmās.

Nemot vērā kompleksa risinājuma sarežģītību, būvniecības nozares vidējie un lielie uzņēmumi šādu automatizāciju uztic ERP (vai PMIS) sistēmu izstrādes uzņēmumiem. Izstrādes uzņēmumi lielajiem klientiem izveido vienotu visaptverošu modulāru sistēmu, lai pārvaldītu daudzus dažādus informācijas slāņus, tostarp materiālu un resursu aprēķinus.



## NODAĻA 5.4.

### BŪVΝIECĪBAS ERP UN PMIS SISTĒMAS

#### Būvniecības ERP -sistēmas uz aprēķinu un tāmju piemēra

Modulārās ERP sistēmas integrē dažādus atribūtu (informācijas) slāņus un datu plūsmas vienā visaptverošā sistēmā, ļaujot projektu vadītājiem sinhronizēti pārvaldīt resursus, finanses, logistiku un citus projekta aspektus vienā platformā. Būvniecības ERP sistēma darbojas kā būvniecības projektu "smadzenes", vienkāršojot atkārtotus procesus ar automatizācijas palīdzību, nodrošinot pārredzamību un kontroli visā būvniecības procesā.

**Būvniecības ERP** -sistēmas (Enterprise Resource Planning) ir visaptveroši programmatūras risinājumi, kas izstrādāti, lai pārvaldītu un optimizētu dažādus būvniecības procesa aspektus. Būvniecības ERP sistēmu pamatā ir izmaksu aprēķina un plānošanas pārvaldības moduļi, kas padara tās par svarīgu rīku efektīvai resursu plānošanai.

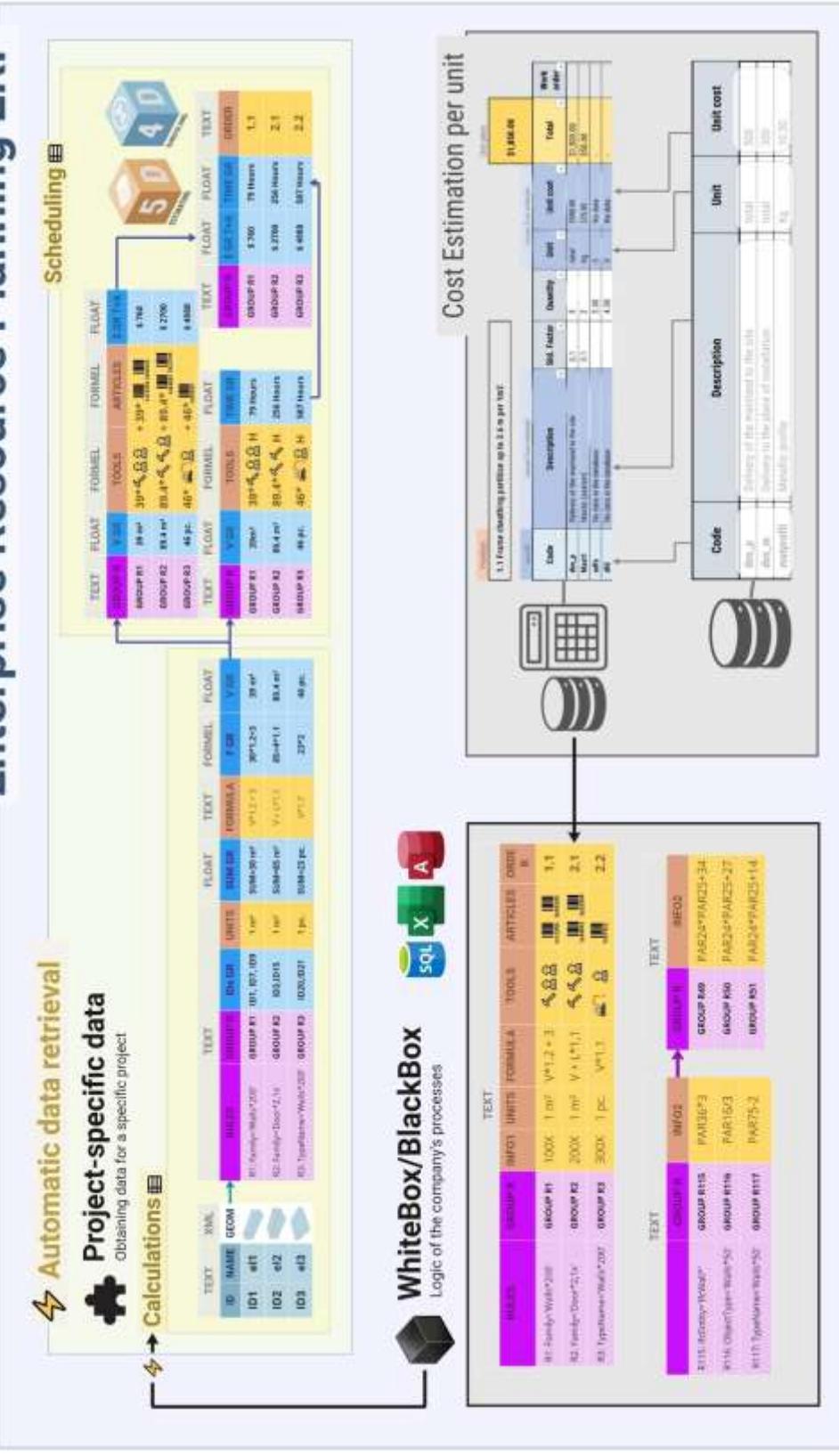
ERP moduļi -sistēmas ļauj lietotājiem strukturētā veidā ievadīt, apstrādāt un analizēt datus, kas aptver dažādus projekta aspektus, tostarp materiālu un darbaspēka izmaksu uzskaiti, iekārtu izmantošanu, loģistikas pārvaldību, cilvēkresursus, kontaktus un citas būvniecības darbības.

Viens no sistēmas funkcionālajiem blokiem ir biznesa loģikas automatizācijas modulis - BlackBox/WhiteBox, kas pilda procesa vadības centra lomu.

BlackBox /WhiteBox ļauj speciālistiem, kas izmanto ERP sistēmu, elastīgi pārvaldīt dažādus uzņēmējdarbības aspektus, kurus jau iepriekš konfigurējuši citi lietotāji vai administratori, izmantojot piekļuves tiesības. ERP sistēmu kontekstā termini *BlackBox* un *WhiteBox* attiecas uz sistēmas iekšējās loģikas pārredzamības un kontrolējamības līmeņiem:

- **BlackBox** ("melnā kaste") - lietotājs mijiedarbojas ar sistēmu, izmantojot saskarni, bez piekļuves procesa izpildes iekšējai loģikai. Sistēma pati veic aprēķinus, pamatojoties uz iepriekš definētiem noteikumiem, kas ir slēpti no galalietotāja. Viņš ievada datus un saņem rezultātu, nezinot, kādi atribūti vai koeficienti tika izmantoti iekšienē.
- **WhiteBox** ("baltā kaste") - procesa loģika ir pieejama apskatei, pielāgošanai un modificēšanai. Uzlabotie lietotāji, administratori vai integratori var manuāli definēt datu apstrādes algoritmus, aprēķinu noteikumus un mijiedarbības scenārijus starp projekta vienībām.

Enterprise Resource Planning ERP



**Attēls 5.4-1 Būvniecības ERP -sistēmas arhitektūra, lai iegūtu tāmes un darbu grafikus, manuāli aizpildot apjoma atrībūtus.**

Piemēram, pieredzējis lietotājs vai administrators nosaka noteikumu: kuri aplēses atribūti jāreizina savā starpā vai jāgrupē pēc noteikta raksturlieluma un kur jāreģistrē galīgais rezultāts. Vēlāk mazāk apmācīti speciālisti, piemēram, tāmēšanas inženieri, vienkārši ielādē jaunus datus ERP, izmantojot lietotāja saskarni, un iegūst gatavas tāmes, grafikus vai specifikācijas, nerakstot kodu vai neizprotot loģikas tehniskās detaļas.

Iepriekšējās nodaļās aprēķinu un loģikas moduļi tika aplūkoti LLM mijiedarbības kontekstā. ERP vidē šādi aprēķini un transformācijas notiek moduļos, kas paslēpti aiz pogu un veidlalu saskarnes.

Turpmākajā piemērā ( 5.4-1. attēls) ERP -sistēmas administrators BlackBox /WhiteBox modulī ir definējis noteikumus, lai saskaņotu aplēšu vienību atribūtus ar QTO grupēšanas atribūtiem. Pateicoties šim BlackBox/WhiteBox moduļa konfigurējumam (ko veicis vadītājs vai administrators), lietotājs (tāmētājs vai inženieris), manuāli pievienojot daudzuma vai apjoma atribūtu, izmantojot ERP lietotāja saskarni, automātiski saņem gatavas tāmes un darba grafikus. Šādā veidā iepriekšējās nodaļās aplūkotie aprēķinu un tāmju ģenerēšanas procesi ar koda palīdzību ERP iekšienē kļūst par daļēji automatizētu konveijeru.

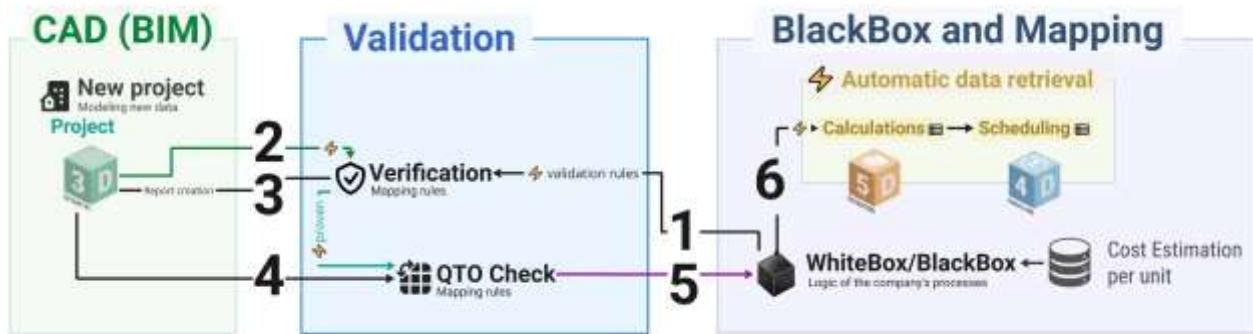
Savienojot šo daļēji automatizēto procesu ar tilpuma atribūtiem no CAD (BIM) modeļiem ( 4.1-13. attēls), piemēram, ielādējot CAD projektu ERP modulī, kas šim nolūkam ir iepriekš konfigurēts, datu plūsma kļūst par sinhronizētu mehānismu, kas spēj autonomi un nekavējoties atjaunināt atsevišķu elementu grupu vai visa projekta vērtību, reaģējot uz jebkādām izmaiņām projektēšanas posmā, ielādējot CAD modeli ERP sistēmā.

Lai izveidotu automatizētu datu plūsmu ( 5.4-2. attēls) starp CAD (BIM) un ERP sistēmām, ir strukturēti jādefinē pamatprocesi un prasības attiecībā uz datiem no CAD (BIM) modeļu datubāzēm, kā tas jau tika aplūkots iepriekšējā nodaļā "Prasības un datu kvalitātes nodrošināšana ". Šis process ERP sistēmā ir sadalīts līdzīgos posmos:

- **Validācijas noteikumu izveide (1)**, kuriem ir svarīga nozīme, lai nodrošinātu datu precizitāti, kas tiek ievadīti ERP sistēmā. Validācijas noteikumi kalpo kā filtri, kas validē vienības un to atribūtus, ļaujot sistēmā ievadīt tikai tos elementus, kas atbilst prasībām. Vairāk par verifikāciju un validāciju var uzzināt nodaļā " prasību izveide un datu kvalitātes validēšana ".
- Pēc tam ERP **notiek verifikācijas process (2)**, kas apstiprina, ka visi projekta vienības elementi ar to atribūtiem un vērtībām ir izveidoti pareizi un ir gatavi nākamajiem apstrādes posmiem.
- Ja ir problēmas ar nepilnīgiem atribūtu datiem, **tieki sagatavots ziņojums (3)**, un projekts kopā ar norādījumiem par labojumiem tiek nosūtīts pārskatīšanai, līdz ir gatavs nākamajai atkārtojumam.
- Kad projekta dati ir apstiprināti un pārbaudīti, tos izmanto citā ERP modulī **(4)**, **lai izveidotu daudzuma ņemšanas tabulas (QTO)**, kas izveido daudzumu atribūtus vienību grupām, materiāliem un resursiem saskaņā ar iepriekš izveidotiem noteikumiem (WhiteBox/BlackBox).
- Datus, kas sagrupēti pēc atbilstības noteikumiem vai QTO, automātiski **integrē ar aprēķiniem**

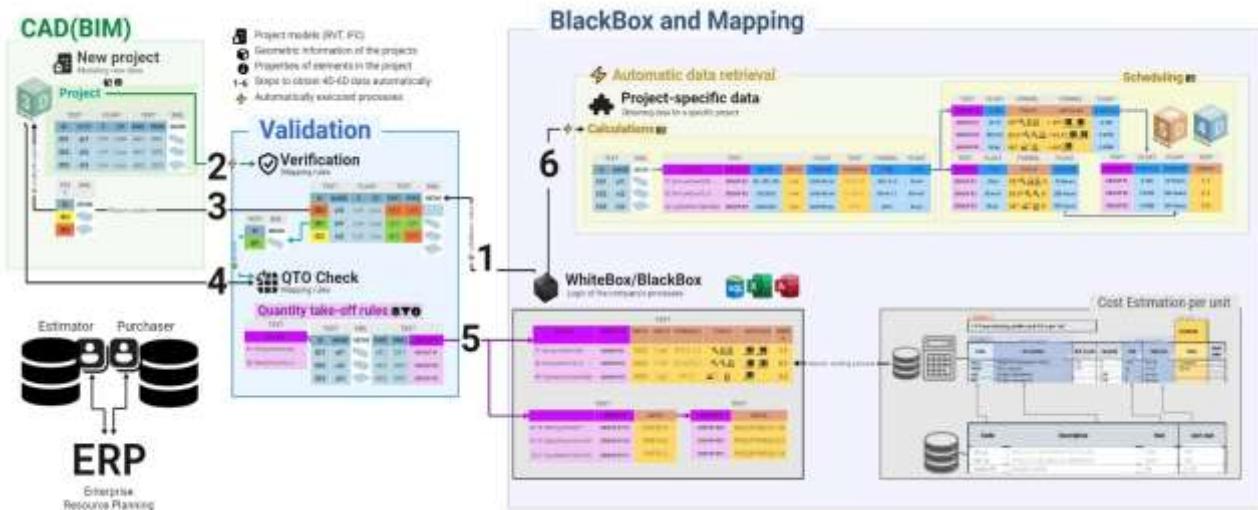
(piemēram, izmaksas un laiks) (5).

- ERP -sistēmas pēdējā posmā lietotājs, reizinot darbības jomas atribūtus no QTO tabulas ar procesa tabulu atribūtiem (piemēram, aplēstajiem posteņiem), **automātiski generē aprēķinu rezultātus (6)** piemēram, izmaksu aplēses, darba grafikus vai CO<sub>2</sub> emisijas) katrai struktūru grupai un projektam kopumā.

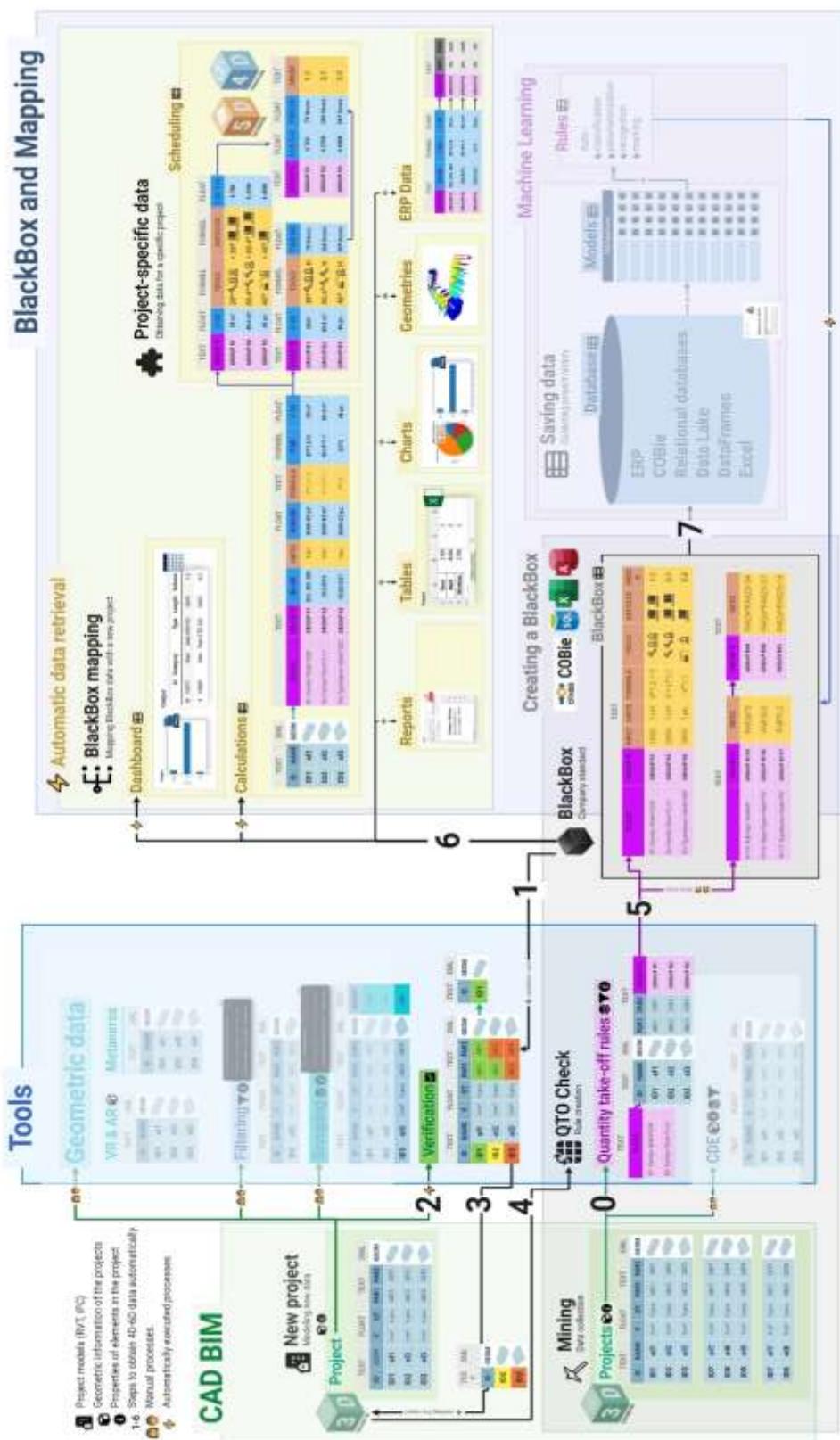


Attēls 5.4-2 Būvniecības ERP -sistēmas ar CAD (BIM) arhitektūra, sākot no apstiprināšanas noteikumu izveides (1) līdz izmaksu un darba grafiku automātiskai aprēķināšanai (5-6).

Modulārajā ERP -sistēmā procesi tiek integrēti, izmantojot programmatūru, kas ietver lietotāja saskarni. Aiz saskarnes atrodas aizmugurējā daļa, kurā strukturētas tabulas apstrādā datus, veicot dažādas vadītāja vai administratora iepriekš konfigurētas darbības. Rezultātā lietotājs, pateicoties iepriekš definētai un pielāgotai automatizācijas loģikai (moduļos BlackBox /WhiteBox), saņem pusautomātiski sagatavotus dokumentus, kas atbilst viņa uzdevumiem



Attēls 5.4-3 ERP -sistēma palīdz vadītājiem un lietotājiem pārvietoties starp speciālistu tabulām, lai generētu jaunus datus.



Attēls 5.4-4 ERP -sistēma ir integrēta ar analītiskajiem rīkiem un automatizē uzņēmuma lēmumu pieņemšanas procesu.

Līdzīgi arī procesi ERP -sistēmā no sākuma līdz galīgajam aprēķinam (1.-6. posms, 5.4-3. attēls) ir savstarpēji saistīti posmu ķēde, kas galu galā nodrošina plānošanas pārredzamību, efektivitāti un precizitāti.

Mūsdienīgās būvniecības ERP -sistēmas ietver ne tikai izmaksu un laika grafiku aprēķināšanas moduļus, bet arī desmitiem citu iepriekš konfigurētu moduļu, kas parasti aptver dokumentu pārvaldību, projekta progresu uzraudzību, līgumu pārvaldību, piegādes ķēdi un logistiku, kā arī integrāciju ar citām biznesa sistēmām un platformām. ERP integrētie analītiskie rīki ļauj lietotājiem automatizēt vadības paneļu izveidi, lai uzraudzītu projekta KPI (KPI - galvenie darbības rādītāji). Tas nodrošina centralizētu un konsekventu visu būvniecības projekta aspektu pārvaldību, mēģinot apvienot lielu skaitu lietojumprogrammu un sistēmu vienā platformā.

Nākotnē ERP -analītika tiks izmantota kombinācijā ar mašīnmācīšanos, lai palielinātu precizitāti un optimizētu turpmāko projekta atribūtu aprēķināšanas procesu. No ERP sistēmām analizētie un savāktie dati un atribūti lielos datos (5.4-4. attēls) nākotnē veidos pamatu prognozēšanas modeļu izveidei, kas spēs precīzi paredzēt iespējamos kavējumus, riskus vai, piemēram, materiālu izmaksu iespējamās izmaiņas.

Kā alternatīvu ERP, būvniecības nozarē bieži izmanto PMIS (Projektu vadības informācijas sistēmu) - projektu vadības sistēmu, kas paredzēta detalizētai uzdevumu kontrolei atsevišķa būvniecības projekta līmenī.

## **PMIS: starpnieks starp ERP un būvlaukumu.**

Atšķirībā no ERP, kas aptver visu uzņēmuma uzņēmējdarbības procesu ķēdi, PMIS koncentrējas uz konkrēta projekta pārvaldību, laika grafiku, budžetu, resursu un dokumentācijas uzraudzību.

PMIS (Project Management Information System) ir būvniecības projektu vadības programmatūra, kas paredzēta visu projekta aspektu plānošanai, izsekošanai, analīzei un pārskatu sagatavošanai.

PMIS ļauj pārvaldīt dokumentus, grafikus, budžetus, un pirmajā brīdī var šķist, ka PMIS ir ERP dublējošs risinājums, taču galvenā atšķirība ir pārvaldības līmenis:

- **ERP** ir vērsta uz uzņēmuma uzņēmējdarbības procesiem kopumā: izmaksu, līgumu, iepirkumu, cilvēkresursu un resursu pārvaldību uzņēmuma līmenī.
- **PMIS** koncentrējas uz individuālu projektu vadību, nodrošinot detalizētu plānošanu, izmaiņu kontroli, pārskatu sniegšanu un dalībnieku koordinēšanu.

Daudzos gadījumos tieši ERP -sistēmām jau ir pietiekama funkcionalitāte, un PMIS ieviešana klūst drīzāk par ērtības un uzņēmuma vēlmju jautājumu. Daudzi darbuzņēmēji un klienti izmanto PMIS nevis tāpēc, ka tas ir nepieciešams, bet gan tāpēc, ka to uzspiež pārdevējs vai liels klients, kurš vēlas apkopot datus konkrētā platformā.

Jāatzīmē, ka starptautiskajā būvniecības projektu vadības terminoloģijā ir arī citi atsevišķi populāri jēdziensi, piemēram, PLM (Product Lifecycle Management) un EPC un EPC-M (Engineering,

Procurement and Construction Management) - līgumu slēgšanas metodes būvniecības nozarē.

Ja uzņēmums jau izmanto ERP ar projektu vadības moduļiem, PMIS ieviešana var būt nevajadzīgs savienojums, kas dublē funkcionalitāti. Tomēr, ja procesi nav automatizēti un dati ir sadrumstaloti, PMIS var būt ērtāks un vieglāk kopjams rīks.

## **Spekulācijas, peļņa, izolētība un trūkums pārredzamība ERP un PMIS**

Neraugoties uz ārējo saskarņu un procedūru vienkāršību, būvniecības ERP un PMIS -sistēmas vairumā gadījumu ir slēgti un neelastīgi risinājumi. Šādas sistēmas parasti tiek piegādātas kā viena ražotāja iepriekš konfigurēta programmatūras pakete ar ierobežotu piekļuvi iekšējām datubāzēm un procesu loģikai.

CAD-(BIM-) piedāvātāji arvien vairāk pārņem šādu sistēmu izstrādi un kontroli, jo to datubāzēs ir ERP sistēmām nepieciešamā informācija: projekta elementu kvantitatīvie un tilpuma atribūti. Tomēr tā vietā, lai nodrošinātu piekļuvi šiem datiem atvērtā vai mašīnlasāmā formātā, pārdevēji piedāvā tikai ierobežotus lietotāju scenārijus un slēgtu apstrādes logiku - iepriekš definētu BlackBox moduļos. Tas samazina sistēmas elastību un neļauj to pielāgot konkrētiem projekta apstākļiem.

Ierobežota datu pārredzamība joprojām ir viens no galvenajiem izaicinājumiem digitālo procesu jomā būvniecībā. Slēgta datu bāzu arhitektūra, piekļuves trūkums pilnām būvelementu atribūtu kopām, koncentrēšanās uz *BlackBox* automatizācijas *moduļiem* un atvērtu saskarņu trūkums ievērojami palielina dokumentu birokrātijas risku. Šādi ierobežojumi rada sastrēgumus lēmumu pieņemšanas procesā, apgrūtina informācijas pārbaudi un paver iespēju datu slēpšanai vai spekulācijām ERP/PMIS sistēmās. Lietotāji parasti saņem tikai ierobežotu piekļuvi - vai tā būtu ierobežota saskarne vai daļēja API - bez iespējas tieši mijiedarbīties ar primārajiem datu avotiem. Tas ir īpaši svarīgi, ja runa ir par parametriem, kas automātiski generēti no CAD projektiem, piemēram, tilpumiem, laukumiem un daudzumiem, ko izmanto QTO aprēķinos.

Tā rezultātā daudzi būvniecības uzņēmumi tā vietā, lai meklētu efektivitāti, izmantojot procesu automatizāciju, atvērtus datus, darījumu izmaksu samazināšanu un jaunu uzņēmējdarbības modeļu izveidi, koncentrējas uz ārējo parametru pārvaldību - manipulāciju ar faktoriem, korekcijas koeficientiem un aprēķinu metodēm, kas ietekmē projekta izmaksas slēgtās ERP/PMIS platformās. Tas rada iespēju spekulācijām, izkroplo reālās ražošanas izmaksas un mazina uzticību starp visiem būvniecības procesa dalībniekiem.

Būvniecībā peļņa veidojas kā starpība starp ieņēmumiem no pabeigta projekta un mainīgajām izmaksām, kas ietver projektēšanas, materiālu, darbaspēka un citas tiešas izmaksas, kurās ir tieši saistītas ar projekta īstenošanu. Tomēr galvenais faktors, kas ietekmē šo izmaksu vērtību, ir ne tikai tehnoloģija vai loģistika, bet arī aprēķinu ātrums un precizitāte, kā arī uzņēmuma vadības lēmumu kvalitāte.

Problēmu saasina fakts, ka lielākajā daļā būvniecības uzņēmumu izmaksu aprēķināšanas procesi ir nepārredzami ne tikai klientiem, bet arī pašiem darbiniekiem, kuri nav saistīti ar tāmēšanas vai finanšu nodalām. Šāda noslēgtība veicina priviliģētas speciālistu grupas veidošanos uzņēmumā - "finanšu zināšanu" nesēju, kuriem ir ekskluzīvas tiesības redīgēt atribūtus un korekcijas koeficientus ERP/PMIS sistēmās. Šie darbinieki kopā ar uzņēmumu vadītājiem var faktiski kontrolēt projekta finanšu loģiku.

Šādos apstākļos tāmētāji kļūst par "finanšu žonglieriem", kas balansē starp uzņēmuma peļņas maksimizēšanu un nepieciešamību saglabāt konkurētspējīgu cenu klientam. Tajā pašā laikā viņiem jāizvairās no acīmredzamām un rupjām manipulācijām, lai neapdraudētu uzņēmuma reputāciju. Tieši šajā posmā tiek noteikti koeficienti, lai slēptu pārvērtētus apjomus vai materiālu un darbu izmaksas.

Rezultātā būvniecības nozarē strādājošo uzņēmumu efektivitātes un rentabilitātes palielināšanas galvenā shēma ir nevis automatizācija un lēmumu pieņemšanas procesu paātrināšana, bet gan spekulācijas ar materiālu un darbu cenām ( 5.4-5. attēls). Darbu un materiālu izmaksu pārvērtēšanu veic pelēkā grāmatvedība slēgtās ERP /PMIS - sistēmāsuzpūšot procentus virs vidējām tirgus cenām materiāliem vai darbu apjomiem ar koeficientu palīdzību ( 5.1-6. att.), kas tika aplūkoti nodaļā "Aprēķinu sastādīšana un darbu izmaksu aprēķināšana, pamatojoties uz resursu bāzi".

Rezultātā klients saņem aprēķinu, kas neatspoguļo reālās izmaksas vai darbu apjomu, bet ir daudzu slēptu iekšējo koeficientu atvasinājums. Tajā pašā laikā apakšuzņēmēji, cenšoties izpildīt ģenerāluzņēmēja noteiktās pārāk zemās cenas, bieži vien ir spiesti iegādāties lētākus un nekvalitatīvus materiālus, kas pasliktina būvniecības galīgo kvalitāti.

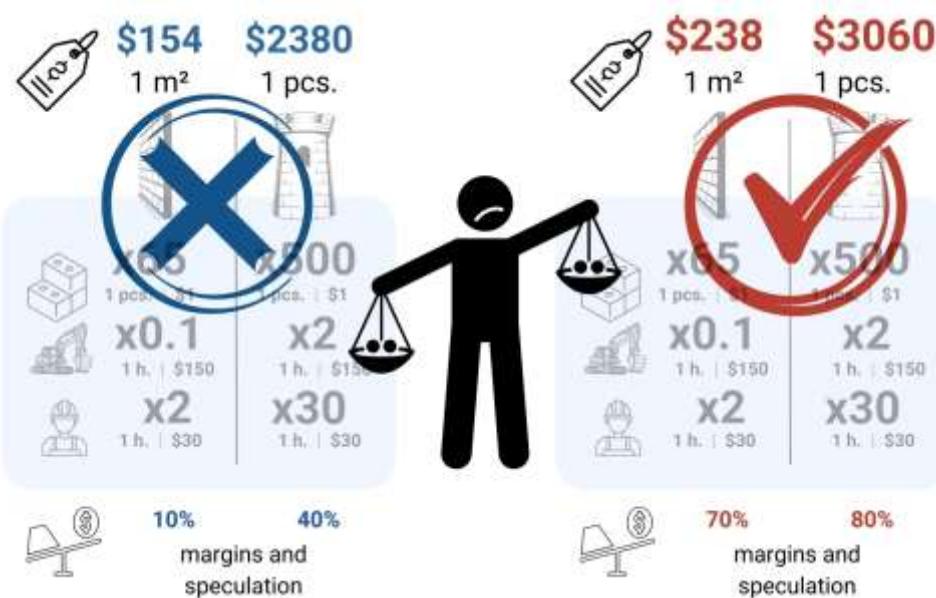
Spekulatīvais process, kurā peļņa tiek meklēta no tukša gaisa, galu galā kaitē gan klientiem, kuri saņem neuzticamus datus, gan izpildītājiem, kuri ir spiesti meklēt arvien jaunus spekulatīvos modeļus

Tādējādi, jo lielāks projekts, jo augstāks birokrātijas līmenis datu un procesu pārvaldībā. Katrā posmā un katrā modulī bieži vien slēpjās necaurspīdīgi koeficienti un piemaksas, kas iestrādātas aprēķinu algoritmos un iekšējās procedūrās. Tas ne tikai apgrūtina revīziju, bet arī ievērojamīzi kroplo projekta finansiālo ainu. Lielos būvniecības projektos šāda prakse bieži vien novēd pie tā, ka galīgās izmaksas pieaug vairākkārtīgi (dažkārt pat desmitkārtīgi), bet reālie apjomi un izmaksas paliek ārpus klienta efektīvas kontroles ( 2.1-3. attēls. lielu infrastruktūras projektu plānoto un faktisko izmaksu salīdzinājums Vācijā.

Saskaņā ar McKinsey & Company 2016. gada ziņojumu *Imagine the Digital Future of Construction* (2016) lielie būvniecības projekti tiek pabeigti vidēji par 20% vēlāk, nekā plānots, un līdz pat 80% pārsniedz budžetu [107].

Aprēķinu un budžeta veidošanas nodaļas kļūst par visaizsargātāko uzņēmuma saikni. Piekļuve tām ir stingri ierobežota pat iekšējiem speciālistiem, un slēgtās loģikas un datubāzu struktūras dēļ nav iespējams objektīvi un bez izkroplojumiem novērtēt projektu lēmumu efektivitāti. Pārredzamības trūkums novēd pie tā, ka uzņēmumi ir spiesti nevis optimizēt procesus, bet cīnīties par izdzīvošanu, "radoši" pārvaldot skaitļus un faktorus ( 5.3-1. attēls, 5.1-6. attēls - piemēram, parametrs "Bid.

Faktors").



Attēlā: 5.4-5 Norēķinu līmeņa spekulācijas koeficienti ir uzņēmumu galvenā peļņa un māksla žonglēt starp darba kvalitāti un reputāciju.

Tas viss liek apšaubīt slēgtu ERP/PMIS sistēmu turpmāku izmantošanu būvniecībā. Nemot vērā digitālo transformāciju un pieaugošās klientu prasības pēc pārredzamības ( 10.2-3. attēls), maz ticams, ka projektu realizācija ilgtermiņā paliks atkarīga no patentētiem risinājumiem, kas ierobežo elastību, kavē integrāciju un kavē uzņēmējdarbības attīstību.

Lai cik izdevīgi būvniecības uzņēmumiem būtu strādāt ar datu silosiem un necaurspīdīgiem datiem slēgtās datubāzēs, būvniecības nozares nākotne neizbēgami būs saistīta ar pāreju uz atvērtām platformām, mašīnlasāmām un pārredzamām datu struktūrām un uz uzticību balstītu automatizāciju. Šī transformācija tiks virzīta no augšas - zem klientu, regulatoru un sabiedrības spiediena, kas arvien vairāk pieprasī atbildību, ilgtspēju, pārredzamību un ekonomisko pamatotību.

## Slēgtas ERP /PMIS ēras beigas: būvniecības nozarei nepieciešamas jaunas pieejas.

Apjomīgu moduļu ERP/PMIS sistēmu izmantošana, kas sastāv no desmitiem miljonu kodu rindu, ārkārtīgi apgrūtina jebkādas izmaiņas tajās. Šādā gadījumā pāreja uz jaunu platformu, ja ir jau iepriekš uzņēmumam konfigurēti moduļi, desmitiem tūkstošu rakstu resursu datubāzēs (5.1-3. attēls) un tūkstošiem gatavu aprēķinu (5.1-6. attēls), kļūst par dārgu un ilgstošu procesu. Jo vairāk koda un mantotās arhitektūras - jo augstāks iekšējās neefektivitātes līmenis, un katrs jauns projekts situāciju tikai pasliktina. Daudzos uzņēmumos datu migrācija un jaunu risinājumu integrācija kļūst par vairāku gadu epopeju, ko pavada nemitīgi pārstrādāšanas darbi un nebeidzami kompromisu meklējumi. Rezultāts bieži vien ir atgriešanās pie vecajām, pazīstamajām platformām, neraugoties uz to ierobežojumiem.

Kā uzsvērts Vācijas Melnās grāmatas ziņojumā [108] par sistēmiskām kļūdām būvniecības datu pārvaldībā, informācijas sadrumstalotība un centralizētas piejas trūkums tās pārvaldībai ir galvenais neefektivitātes iemesls. Bez standartizācijas un integrācijas dati zaudē savu vērtību, kļūstot drīzāk par arhīvu, nevis par pārvaldības rīku.

Viens no galvenajiem datu kvalitātes zudumu iemesliem ir neatbilstoša būvniecības projektu plānošana un kontrole, kas bieži vien izraisa ievērojamu izmaksu pieaugumu. Melnās grāmatas sadaļā "Uzmanības centrā: izmaksu eksplozija" analizēti galvenie faktori, kas veicina šīs nevēlamās sekas. Tie ietver neatbilstošu vajadzību analīzi, priekšizpētes trūkumu un nesaskaņotu plānošanu, kas rada papildu izmaksas, no kurām varēja izvairīties.

Nobriedušā uzņēmuma IT ekosistēmā novecojušas sistēmas nomaiņa ir salīdzināma ar nesošās kolonnas nomaiņu jau uzceltā ēkā. Nepietiek vienkārši noņemt veco un uzstādīt jaunu - ir svarīgi to izdarīt tā, lai ēka paliktu stabila, griesti nesabruktu un visas komunikācijas turpinātu darboties. Tieši šeit slēpjas grūtības: jebkura kļūda var radīt nopietnas sekas visai uzņēmuma sistēmai.

Tomēr lielo būvniecības nozarei paredzēto ERP produktu izstrādātāji joprojām izmanto uzrakstītā koda apjomu kā argumentu par labu savai platformai. Specializētās konferencēs joprojām var dzirdēt tādas frāzes kā: "Šādas sistēmas atjaunošana prasītu 150 cilvēkgadus", lai gan lielākā daļa šādu sistēmu funkcionalitātes slēpj datu bāzes un diezgan vienkāršas funkcijas darbam ar tabulām, kas ievietotas īpašā fiksētā lietotāja saskarnē. Praksē "150 cilvēkgadu" koda apjoms kļūst par slogu, nevis par konkurences priekšrocību. Jo vairāk koda - jo lielākas uzturēšanas izmaksas, jo grūtāk pielāgoties jauniem apstākļiem un jo augstāks ienākšanas slieksnis jauniem izstrādātājiem un klientiem.

Daudzas moduļu ēku sistēmas mūsdienās atgādina apgrūtinošas un novecojušas "Frankenšteina konstrukcijas", kurās jebkuras neuzmanīgas izmaiņas var novest pie neveiksmēm. Katrs jauns modulis sarežģī jau tā pārslogoto sistēmu, padarot to par labirintu, kurā orientējas tikai daži speciālisti, tādējādi vēl vairāk apgrūtinot tās uzturēšanu un modernizāciju.

Sarežģītību apzinās arī paši izstrādātāji, kuri periodiski pārtrauc refaktorizāciju - arhitektūras pārskatīšanu, lai ņemtu vērā jauno tehnoloģiju rašanos. Tomēr, pat ja refaktorizācija tiek veikta regulāri, sarežģītība neizbēgami pieaug. Šādu sistēmu arhitekti pierod pie augošās sarežģītības, bet jauniem lietotājiem un speciālistiem tā kļūst par nepārvaramu šķērsli. Rezultātā visas zināšanas koncentrējas dažu izstrādātāju rokās, un sistēma vairs nav mērogojama. Īstermiņā šādi eksperti ir noderīgi, bet ilgtermiņā viņi kļūst par daļu no problēmas.

*Organizācijas turpinās integrēt "mazos" datus ar lielajiem datiem, un ir muļķīgi uzskatīt, ka viena lietojumprogramma, lai cik dārga vai spēcīga tā būtu, var apstrādāt visu [109].*

- Phil Simon, Conversations About Collaboration podkāsta vadītājs.

Rodas pamatots jautājums: vai mums tiešām ir vajadzīgas tik apgrūtinošas un slēgtas sistēmas, lai aprēķinātu darbu izmaksas un termiņus tabulu veidā, ja citas nozares jau sen spēj tikt galā ar līdzīgiem uzdevumiem, izmantojot analītiskus rīkus ar atvērtiem datiem un pārredzamu loģiku?

Pašlaik būvniecības nozarē joprojām ir pieprasītas slēgtas moduļu platformas, galvenokārt izmaksu uzskaites specifikas dēļ (5.1-7. attēls). Šādas sistēmas bieži tiek izmantotas pelēko vai necaurspīdīgo shēmu vadīšanai, kas ļauj slēpt reālās izmaksas no pasūtītāja. Tomēr, nozarei digitāli nobriestot, galvenokārt klientiem, un pārejot uz tā saukto "Uberizēto laikmetu", starpnieki, proti, būvniecības uzņēmumi ar savām ERP, zaudēs savu nozīmi laika un izmaksu aprēķinos. Tas uz visiem laikiem mainīs būvniecības nozares seju. Vairāk lasiet grāmatas pēdējā daļā un nodaļā "Būvniecība 5.0: Kā pelnīt naudu, kad vairs nevar paslēpties".

Tūkstošiem pēdējo 30 gadu laikā uzkrāto mantoto risinājumu, kuru izstrādē ieguldīti tūkstošiem cilvēkgadu, sāks strauji izzust. Pāreja uz atvērtu, pārredzamu un elastīgu datu pārvaldību ir neizbēgama. Jautājums ir tikai par to, kuri uzņēmumi spēs pielāgoties šīm pārmaiņām un kuri paliks vecā modeļa ķīlnieki.

Līdzīga situācija ir vērojama CAD (BIM) rīku jomā, kuru dati šodien aizpilda projektēšanas vienību tilpuma parametrus ERP/PMIS sistēmās. Sākotnēji BIM ideja (izstrādāta vēl 2002. gadā [110]) bija balstīta uz vienotas integrētas datubāzes koncepciju, bet praksē šodien darbam ar BIM ir nepieciešams vesels specializētas programmatūras un formātu kopums. Tas, kam bija paredzēts vienkāršot projektēšanas un būvniecības pārvaldību, ir pārtapis par vēl vienu patentētu risinājumu slāni, kas sarežģī integrāciju un samazina uzņēmējdarbības elastību.

## Nākamie soļi: efektīva projektu datu izmantošana

Šajā daļā mēs parādījām, kā strukturēti dati klūst par pamatu precīziem būvniecības projektu izmaksu un laika grafiku aprēķiniem. Automatizējot QTO, grafiku sastādišanas un tāmju veidošanas procesus, tiek samazinātas darbaspēka izmaksas un ievērojami uzlabota rezultātu precizitāte.

Apkopojoš šo daļu, ir vērts izcelt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās piejas ikdienas uzdevumos. Šīs piejas ir universālas - tās ir noderīgas gan uzņēmuma digitālajai transformācijai, gan aprēķinos iesaistīto speciālistu ikdienas darbā:

- Automatizēt ikdienas aprēķinus
  - Mēģiniet atrast standarta darba izmaksu aprēķinus, ar kuriem varat saistīt savu darbu.
  - Analizējiet, kādas metodes tiek izmantotas, lai aprēķinātu izmaksas vai aprēķinātu būvlaukuma darbus vai procesus jūsu valstī (5.1-7. attēls).
  - Ja strādājat ar CAD sistēmu - apgūstiet specifikācijas un QTO datu automātiskās ieguves funkcijas savā CAD (BIM-) programmatūrā.
  - Izmantojet LLM, lai rakstītu projektu kodu aprēķinu automatizēšanai.
- Izstrādājiet savus rīkus QTO

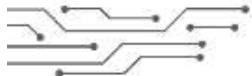
- Izveidojiet skriptus vai tabulas, lai automatizētu apjoma skaitīšanu
- Standartizēt elementu kategorijas un grupas, lai nodrošinātu konsekventu pieeju novērtēšanai.
- dokumentēt aprēķinu metodoloģiju, lai nodrošinātu rezultātu atkārtojamību jaunos projektos.

■ Integrēt dažādus projekta aspektus savā darbā.

- Ja strādājat ar modulārām sistēmām, mēģiniet vizualizēt procesus ne tikai diagrammās vai diagrammās, bet arī datu līmenī, īpaši tabulās.
- Apgūstiet no CAD datubāzēm iegūto datu automātisku apvienošanu ar aprēķiniem - ar Python kodu, izmantojot grupēšanu, filtrēšanu un apkopošanu.
- Izveidot skaidras QTO grupu vizualizācijas, lai kolēgiem un klientiem sniegtu sarežģītu informāciju.

Šie pasākumi palīdzēs izveidot ilgtspējīgu aprēķinu sistēmu, kas balstīta uz automatizāciju un datu standartizāciju. Šāda pieeja uzlabos precīzitāti un samazinās ikdienas aprēķinu jautājumu rutīnu.

Turpmākajās nodaļās uzmanība pievērsta CAD - (BIM-) produktu tehniskajiem aspektiem un iemesliem, kāpēc CAD datubāzes joprojām ir grūti integrēt uzņēmumu uzņēmējdarbības procesos. Ja jūs šobrīd neinteresē BIM ieviešanas būvniecībā vēsture, CAD rīku attīstība un darba ar šīm tehnoloģijām tehniskie aspekti, varat doties uzreiz uz grāmatas septīto daļu "Uz datiem balstīta lēmumu pieņemšana".



## MAKSIMĀLI ĒRTA DRUKĀTĀ VERSIJA

Jūsu rīcībā ir bezmaksas digitālā versija **Data-Driven Construction**. Ērtākam darbam un ātrākai piekļuvei materiāliem iesakām pievērst uzmanību **drukātajam izdevumam**:



**Vienmēr pa rokai:** grāmata drukātā formātā kļūs par uzticamu darba rīku, kas ļaus ātri atrast un izmantot vajadzīgās vizualizācijas un diagrammas jebkurā darba situācijā.

**Augstas kvalitātes ilustrācijas:** visi attēli un grafikas drukātajā izdevumā ir maksimāli kvalitatīvi.

**Ātra piekļuve informācijai:** vienkārša navigācija, iespēja veikt piezīmes, grāmatzīmes un strādāt ar grāmatu jebkurā vietā.

Iegādājoties pilnu drukāto grāmatas versiju, jūs iegūstat ērtu rīku ērtam un efektīvam darbam ar informāciju: iespēju ātri izmantot vizuālos materiālus ikdienas uzdevumos, ātri atrast nepieciešamās shēmas un veikt piezīmes. Turklāt ar savu pirkumu atbalstāt atvērto zināšanu izplatīšanu.

Grāmatas drukāto versiju var pasūtīt: [datadrivenconstruction.io/books](http://datadrivenconstruction.io/books)



## VI DAĻA

# CAD UN BIM: MĀRKETINGS, REALITĀTE UN PROJEKTĒŠANAS DATU NĀKOTNE BŪVΝIECĪBĀ

Grāmatas sestajā daļā kritiski analizēta CAD un BIM -tehnoloģiju attīstība un to ietekme uz datu pārvaldības procesiem būvniecībā. Tajā ir izsekotas BIM koncepcijas vēsturiskās pārvērtības no sākotnējās integrētas datu bāzes idejas līdz pašreizējām mārketinga konstrukcijām, ko veicina programmatūras pārdevēji. Tieks novērtēta patentētu formātu un slēgtu sistēmu ietekme uz projekta datu apstrādes efektivitāti un būvniecības nozares vispārējo darbību. Šīki analizētas dažādu CAD sistēmu savietojamības problēmas un grūtības, kas saistītas ar to integrāciju būvniecības uzņēmumu uzņēmējdarbības procesos. Apskatītas pašreizējās tendences vienkāršotu atvērtu datu formātu, piemēram, USD, virzienā un to potenciālā ietekme uz nozari. Tieks izklāstītas alternatīvas pieejas informācijas iegūšanai no slēgtām sistēmām, tostarp reversās inženierijas metodes. Tieks analizētas mākslīgā intelekta un mašīnmācīšanās izmantošanas perspektīvas, lai automatizētu projektēšanas un datu analīzes procesus būvniecībā. Tieks formulētas prognozes par projektēšanas tehnoloģiju attīstību, kas orientētas uz lietotāju reālajām vajadzībām, nevis programmatūras ražotāju interesēm.

## NODĀLA 6.1.

### BIM- KONCEPCIJU PARĀDĪŠANĀS BŪVΝIECĪBAS NOZARĒ

Sākotnēji šī sestā daļa, kas veltīta CAD (BIM), nebija iekļauta grāmatas pirmajā versijā. Tēmas par patentētiem formātiem, ģeometriskiem kodoliem un slēgtām sistēmām ir pārāk tehniskas, pārslogotas ar detaļām un šķietami nelietderīgas tiem, kas vienkārši vēlas saprast, kā strādāt ar datiem. Tomēr atsauksmes un lūgumi papildināt grāmatas pirmo versiju ar skaidrojumiem parādīja, ka, neizprotot CAD sistēmu iekšējās darbības sarežģītību, ģeometriskos kodolus, formātu daudzveidību un nesaderīgas glabāšanas shēmas vieniem un tiem pašiem datiem, nav iespējams patiesi saprast, kāpēc ražotāju popularizētās koncepcijas bieži vien apgrūtina darbu ar informāciju un kavē pāreju uz atvērto parametrisko projektēšanu. Tāpēc šī daļa grāmatas struktūrā ir ieņēmusi atsevišķu vietu. Ja CAD (BIM) nav jūsu prioritāte, varat pāriet uzreiz uz nākamo daļu - "VII DALĀ: Uz datiem balstīta lēmumu pieņemšana, analītika, automatizācija un mašīnmācīšanās".

### **BIM un atvērtā BIM kā CAD- piegādātāju mārketinga koncepciju rašanās vēsture.**

Līdz ar digitālo datu parādīšanos 20. gadsimta 90. gados datortehnoloģijas tika ieviestas ne tikai uzņēmējdarbības procesos, bet arī projektēšanas procesos, radot tādus jēdzienus kā CAD (datorizētās projektēšanas sistēmas) un vēlāk BIM (būves informācijas modelēšana)

Tomēr, tāpat kā jebkura inovācija, tie nav izstrādes beigu punkts. Tādas koncepcijas kā BIM ir kļuvušas par svarīgu pagrieziena punktu būvniecības nozares vēsturē, taču agri vai vēlu tās var piekāpties labākiem rīkiem un pieejām, kas labāk risinās nākotnes izaicinājumus.

BIM koncepcija, kas parādījās 2002. gadā un ko pārspēja CAD piegādātāju ietekme un apgrūtināja tās ieviešanas sarežģītība, iespējams, nesagaidīs savu trīsdesmito gadadienu, līdzīgi kā rokzvaigzne, kas spoži uzmirdzēja, bet ātri izgaisa. Iemesls ir vienkāršs: datu zinātnieku prasības mainās ātrāk, nekā CAD ražotāji spēj tām pielāgoties.

Saskaroties ar kvalitatīvu datu trūkumu, mūsdienu būvniecības nozares profesionāli pieprasī starplattformu sadarbību un piekļuvi atvērtiem datiem no CAD- projektiem, lai vienkāršotu to analīzi un apstrādi. CAD datu sarežģītība un jucekļīgā CAD datu apstrāde negatīvi ietekmē visus būvniecības procesā iesaistītos: projektētājus, projektu vadītājus, būvlaukuma strādniekus un galu galā arī pasūtītāju.

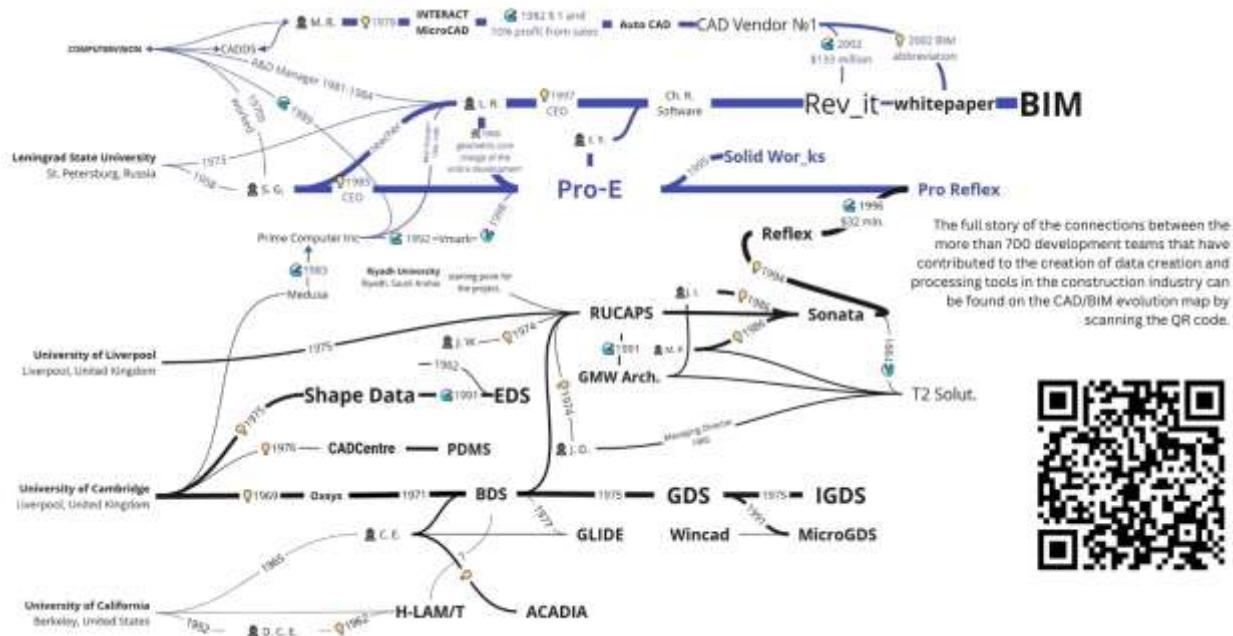
Pilnvērtīgas datu kopas vietā, kas paredzēta ekspluatācijai, klients un ieguldītājs šodien saņem konteinerus CAD- formātos, kuru apstrādei nepieciešami sarežģīti ģeometriskie kodoli, izpratne par datu shēmām, katru gadu atjaunināta API dokumentācija un specializēta CAD programmatūra (BIM), lai strādātu ar šiem datiem. Tajā pašā laikā liela daļa projektēšanas datu paliek neizmantoti.

Mūsdienu projektēšanas un būvniecības pasaule CAD datu piekļuves sarežģītības dēļ projektu vadība ir pārāk sarežģīta. Vidējie un lielie uzņēmumi, kas strādā ar CAD datiem vai izstrādā BIM -risinājumus,

ir spiesti vai nu uzturēt ciešas attiecības ar CAD piegādātāju risinājumiem, lai piekļūtu datiem, izmantojot API, vai arī apiet CAD piegādātāju ierobežojumus, izmantojot dārgus SDK konvertētājus, lai veiktu reverso inženieriju, lai iegūtu atvērtos datus [75].

Patentētu datu pieeja ir novecojusi un vairs neatbilst mūsdieniņu digitālās vides prasībām. Nākotnē uzņēmumi sadalīsies divos veidos: tie, kas efektīvi izmantos atvērtos datus, un tie, kas atstās tirgu.

BIM koncepcija (Building Information Modeling) būvniecības nozarē parādījās līdz ar viena no lielākajiem CAD ražotāju publikāciju - Whitepaper BIM [54] 2002. gadā un, papildināta ar mašīnbūves koncepciju BOM (Bills of Materials), radās no parametriskās pieejas projekta datu izveidei un apstrādei ( 6.1-1. attēls). Parametriskā pieeja projekta datu izveidei un apstrādei bija viena no pirmajām, kas tika ieviesta mašīnbūves projektēšanas sistēmā Pro-E (MCAD). Šī sistēma kļuva par prototipu [111] daudziem mūsdienīgiem CADstostarp tiem, ko šodien izmanto būvniecības nozarē.



Attēls 6.1-1 BIM koncepcijas vēstures karte un līdzīgas koncepcijas.

Žurnālisti un AEC konsultanti, kas līdz 2000. gadu sākumam popularizēja CAD rīkus -pārdevēji, no 2002. gada pievērsās Whitepaper BIM. Tieši 2002.-2004. gadā publicētajam BIM Whitepaper un 2002., 2003., 2005. un 2007. gadā publicētajiem rakstiem bija galvenā loma BIM koncepcijas popularizēšanā būvniecības nozarē [112].

*Būvniecības informācijas modelēšana ir stratēģija..... [CAD piegādātāja uzņēmuma nosaukums], lai pielietotu informācijas tehnoloģijas būvniecības nozarē.*

- BIM Baltā grāmata, 2002 [60]

Līdz 2000. gadu vidum "pētnieki" sāka saistīt BIM-koncepciju, ko 2002. gadā publicēja CAD-pārdevējs, ar agrākiem zinātniskiem darbiem, piemēram, Čārlza Īstmena BDS, kas kļuva par pamatu tādām sistēmām kā GLIDE, GBM, BPM, RUCAPS. Čārlzs Īstmans savā celmlaužā darbā Building Description System (1974) lika teorētiskos pamatus mūsdienu informācijas modelēšanai. Terminus "datubāze" viņa darbā parādās 43 reizes (6.1-2. attēls) - biežāk nekā jebkurš cits, izņemot vārdu "ēka".

Īstmena galvenā ideja bija, ka visa informācija par ēku - no ģeometrijas līdz elementu īpašībām un to savstarpējām attiecībām - ir jāuzglabā vienotā strukturētā datubāzē. No šīs datubāzes var automātiski generēt un analizēt rasējumus, specifikācijas, aprēķinus un atbilstību kodeksiem. Eastman nepārprotami kritizēja rasējumus kā novecojušu un lieku saziņas metodi, norādot uz informācijas dublēšanos, atjaunināšanas problēmām un nepieciešamību veikt manuālus atjauninājumus, kad tiek veiktas izmaiņas. Tā vietā viņš ierosināja izveidot vienotu digitālo modeli datubāzē, kurā visas izmaiņas tiek veiktas vienu reizi un automātiski atspoguļotas visos skatījumos.

Jāatzīmē, ka Īstmens savā koncepcijā vizualizāciju nav izvirzījis centrā. Viņa sistēmas centrā bija informācija: parametri, attiecības, atribūti, analīzes un automatizācijas iespējas. Zīmējumi viņa izpratnē bija tikai viena no datu bāzes datu attēlošanas formām, nevis galvenais projektēšanas informācijas avots.

Pirmajā BIM Baltajā grāmatā, ko publicēja vadošais CAD ražotājs, frāze "datu bāze" tika lietota tikpat bieži kā Čārlza Īstmena BDS - 23 reizes [60] septiņās lappusēs un bija viens no populārākajiem vārdiem dokumentā pēc vārdiem "ēka", "informācija", "modelēšana" un "dizains". Tomēr līdz 2003. gadam termins "datubāze" līdzīgos dokumentos parādās tikai divas reizes [61], un līdz 2000. gadu beigām datubāzu tēma praktiski bija izzudusi no diskusijām par projektēšanas datiem. Rezultātā koncepcija par "vienotu integrētu datu bāzi vizuālai un kvantitatīvai analīzei" tā arī netika pilnībā īstenota.

Tādējādi būvniecības nozare no Čārlza Īstmena (Charles Eastman) progresīvās BDS koncepcijas ar uzsvaru uz datubāzēm un Samuela Geisberga (Samuel Geisberg) idejām par automātisku projektēšanas datu atjaunināšanu no datubāzēm mašīnbūves produktā Pro-E (mūsdienās būvniecībā izmantoto populāro CAD risinājumu priekštece) ir nonākusi līdz pašreiz tirgojamajam BIM, kurā datu pārvaldība, izmantojot datubāzes, gandrīz neminēta, lai gan tieši šāda koncepcija bija sākotnējā teorētiskā pamatā.



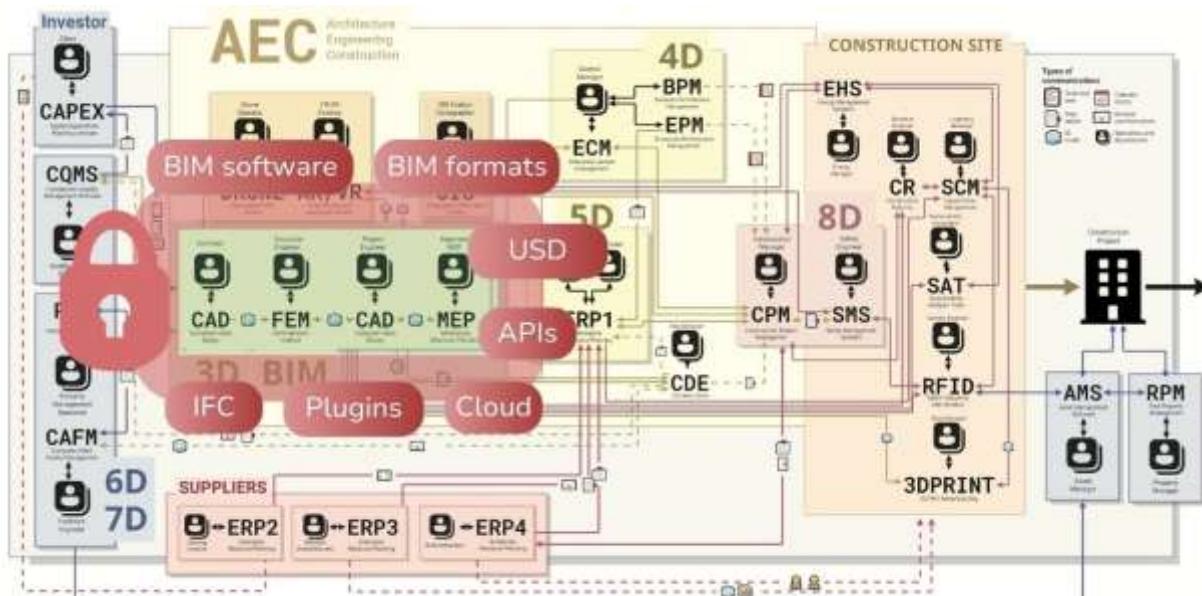
patentētas API.

- **Dati pieder pārdevējiem, nevis lietotājiem.** Projekta informācija ir ieslēgta patentētos formātos vai mākonpakkalpojumos, nevis pieejama atvērtos un neatkarīgos formātos.

Projektēšanas inženieriem un projektu vadītājiem bieži vien nav piekļuves ne CAD datubāzes sistēmām, ne arī formātam, kurā tiek glabāti viņu pašu projekta dati. Tas neļauj ātri pārbaudīt informāciju vai formulēt prasības datu struktūrai un kvalitātei ( 6.1-3. attēls). Lai piekļūtu šādiem datiem, ir nepieciešams vesels specializētu programmu kopums, kas savienots ar API un spraudņiem, kas noved pie pārmērīgas būvniecības nozares procesu birokratizācijas. Tajā pašā laikā šos datus vienlaikus izmanto desmitiem informācijas sistēmu un simtiem speciālistu.

*Mums ir jāspēj pārvadīt visus šos datus [CAD (BIM)], uzglabāt tos digitāli un pārdot dzīves cikla un procesu pārvadības programmatūru, jo **uz katru inženieri [projektētāju], kas kaut ko izveido [CAD programmatūrā], ir desmit cilvēki, kuri strādā ar šiem datiem**" [41].*

- CAD - ražotāja, kas radīja BIM koncepciju, 2005. gadā.



Attēlā 6.1-3 CAD- (BIM-) datu bāzes joprojām ir viena no pēdējām slēgtajām sistēmām IT nodalām un datu pārvadītājiem būvniecības biznesa ekosistēmā.

Kad kļūst skaidrs, ka BIM ir drīzāk datu bāzu komercializācijas līdzeklis, nevis pilnvērtīgs datu bāzu pārvadības rīks, rodas loģisks jautājums: kā atgūt datu kontroli? Atbilde ir izmantot atvērtas datu struktūras, kurās lietotājs, nevis programmatūras ražotājs, kļūst par informācijas īpašnieku.

Būvniecības nozares lietotāji un risinājumu izstrādātāji, tāpat kā viņu kolēģi citās nozarēs, neizbēgami atkāpsies no neskaidrās programmatūras piegādātāju terminoloģijas, kas dominēja pēdējos 30 gados, un pievērsīsies digitalizācijas galvenajiem aspektiem - "datiem" un "procesiem".

Pagājušā gadsimta astoņdesmito gadu beigās par galveno digitālās attīstības jomu būvniecībā tika uzskatīta datu piekļuve un projektu informācijas pārvaldība. Tomēr laika gaitā uzmanība ir mainījusies. Tā vietā, lai izstrādātu pārskatāmas un pieejamas pīejas darbam ar datiem, aktīvi tika popularizēts IFC formāts un atvērtā BIM koncepcija, kas bija mēģinājums novērst speciālistu uzmanību no projektu datu bāzu pārvaldības tēmām.

## Atvērtā formāta IFC parādīšanās būvniecības nozarē

Tā sauktais atvērtais formāts IFC (Industry Foundation Classes) ir standarts, kas nodrošina dažādu CAD (BIM -) sistēmu savstarpēju savietojamību. Tā izstrāde tika veikta organizāciju ietvaros, ko izveidoja un kontrolēja lielākie CAD ražotāji. Pamatojoties uz IFC formātu, divi CAD-uzņēmumi 2012. gadā izstrādāja mārketinga koncepciju OPEN BIM[63].

**IFC** (Industry Foundation Classes) ir atvērts standarts datu apmaiņai būvniecības nozarē, kas izstrādāts, lai nodrošinātu savietojamību starp dažādām CAD - (BIM-) sistēmām.

Atvērtā BIM - koncepcija ietver darbu ar informāciju no CAD datubāzēm un informācijas apmaiņu starp sistēmām, izmantojot atvērtu CAD datu apmaiņas formātu - IFC.

*Atvērtā BIM programma ir mārketinga kampaņa, ko uzsāka... [1 CAD piegādātājs],... [2 CAD piegādātājs] un citi uzņēmumi, lai veicinātu un atvieglotu OPEN BIM koncepcijas koordinētu popularizēšanu visā AEC nozarē, nodrošinot programmas dalībniekiem konsekventu komunikāciju un kopīgu zīmolu.*

- No CAD piegādātāju tīmekļa vietnes, OPEN BIM programma, 2012. gads [113].

IFC no mašīnbūves formāta STEP astoņdesmito gadu beigās pielāgoja Minhenes Tehniskā universitātē, un vēlāk to reģistrēja liels projektēšanas uzņēmums un liels CAD- piegādātājs, 1994. gadā izveidojot IAI (Industry Alliance for Interoperability) [114] ( 6.1-4. attēls). IFC formāts tika izstrādāts, lai nodrošinātu savietojamību starp dažādām CAD sistēmām, un tā pamatā bija mašīnbūves formāta STEP noteiktie principi, kas savukārt radās no IGES formāta, kuru 1979. gadā ar NIST (The National Institute of Standards and Technology) un ASV Aizsardzības departamenta atbalstu izveidoja CAD lietotāju un ražotāju grupa [115].

Tomēr IFC sarežģītā struktūra, tā ciešā atkarība no ģeometriskā kodola, kā arī atšķirības formāta ieviešanā dažādos programmatūras risinājumos ir radījušas daudzas problēmas tā praktiskajā izmantošanā. Ar līdzīgām grūtībām - detalizācijas zudumiem, precizitātes ierobežojumiem un nepieciešamību izmantot starpformātus - iepriekš saskārās mašīnbūves speciālisti, strādājot ar IGES un STEP formātiem, no kuriem radās IFC.



**Attēls 6.1-4 Attīstības komandu un CAD produktu saikņu karte (BIM) [116].**

2000. gadā tas pats CAD - piegādātājs, kas reģistrēja IFC formātu un izveidoja IAI (vēlāk bS) organizāciju, publicē Balto grāmatu "Integrated Design and Manufacturing: Benefits and Rationale". [65]. Dokumentā uzsvērts, cik svarīgi ir saglabāt pilnīgu datu granularitāti, apmainoties ar datiem starp vienas sistēmas programmām, neizmantojot neutrālus formātus, piemēram, IGES, STEP [identiski IFC]. Tā vietā tika ierosināts, ka lietojumprogrammām jābūt tiešai piekļuvei pamatā esošajai CAD datubāzei, lai novērstu informācijas precizitātes zudumus.

2002. gadā tas pats CAD pārdevējs iegādājas parametrisko BOM produktu (3.1-18. attēls, sīkāka informācija trešajā daļā) un uz tā pamata veido BIM koncepciju. Rezultātā būvprojektu datu apmaiņā tiek izmantoti tikai slēgtie CAD formāti vai IFC formāts (STEP), par kura ierobežojumiem 2000. gadā rakstīja pats CAD pārdevējs, kas šo formātu ieviesa būvniecības nozarē.

Sīkāka vairāk nekā 700 būvdatu izveides un apstrādes rīku izstrādes komandu mijiedarbības vēsture ir atspoguļota kartē "CAD (BIM) evolūcija" [116]. [116].

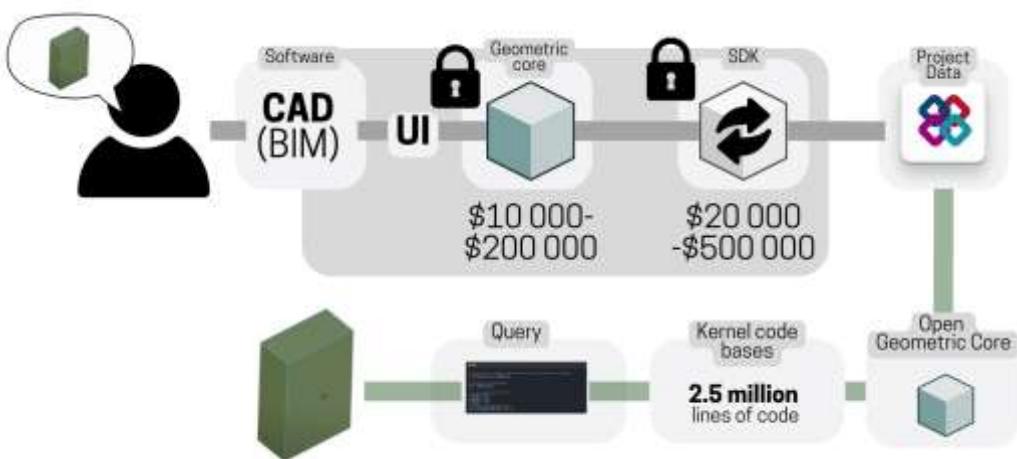
Atvērtā formā IFC sastāv no konstrukcijas elementu ģeometriskā apraksta un metainformācijas apraksta. Ģeometrijas attēlošanai IFC formātā tiek izmantotas dažādas metodes, piemēram, CSG un Swept Solids: tomēr parametriskā attēlošana BREP ir kļuvusi par vadošo standartu elementu ģeometrijas pārnesei IFC formātā, jo šis formāts tiek atbalstīts, eksportējot no CAD- (BIM-)

programmām, un ļauj potenciāli rediģēt elementus, importējot IFC atpakaļ uz CAD programmām.

## IFC formāta problēma atkarībā no ģeometriskā kodola

Vairumā gadījumu, kad IFC ģeometrija ir definēta parametriski (BREP), kļūst neiespējami vizualizēt vai iegūt ģeometriskās īpašības, piemēram, projekta vienību tilpumu vai laukumu, izmantojot tikai IFC failu, jo, lai strādātu ar ģeometriju un vizualizētu to, šajā gadījumā ir nepieciešams ģeometrijas kodols ( 6.1-5. attēls), kura sākotnēji nav.

**Ģeometrijas kodols** ir programmatūras komponents, kas nodrošina pamata algoritmus ģeometrisko objektu izveidei, rediģēšanai un analīzei CAD (CAD), BIM un citās inženierprogrammās. Tas ir atbildīgs par 2D un 3D -ģeometrijas veidošanu un darbībām ar to, piemēram, Bula operācijām, izlīdzināšanu, krustojumiem, transformācijām un vizualizāciju.

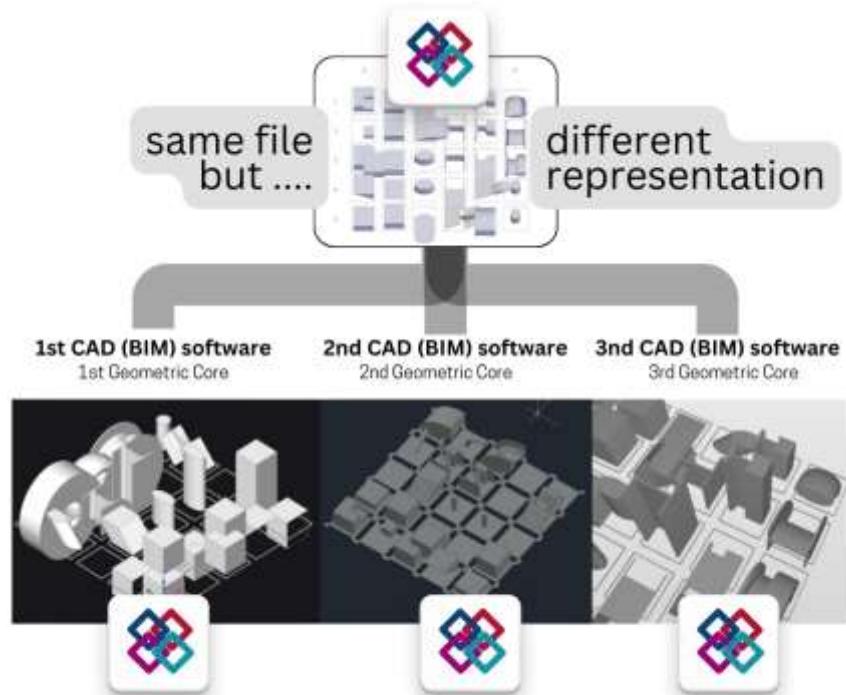


Attēls. 6.1-5 Ģeometrijas izveide, izmantojot CAD- programmatūru, mūsdienās notiek, izmantojot patentētus ģeometrijas kodonus un SDK, kas bieži vien nepieder CAD ražotājiem.

Katrai CAD programmai un visām programmām, kas strādā ar parametriskiem vai IFC formātiem, ir savs vai iegādāts ģeometriskais kodols. Un, ja ar primitīvajiem elementiem IFC -BREP formātā nevar būt problēmu, un programmās ar dažādiem ģeometriskiem kodoliem šos elementus var attēlot līdzīgi, bet bez problēmām ar dažādiem ģeometrisko kodolu dzinējiem, ir pietiekami daudz elementu, kuriem ir savas īpatnības pareizai attēlošanai. Šī problēma ir detalizēti aplūkota starptautiskajā pētījumā " A reference study of IFC software support", kas publicēts 2019. gadā [117].

Tās pašas standartizētās datu kopas sniedz pretrunīgus rezultātus, un ir konstatēti tikai daži kopīgi modeļi, un ir konstatētas nopietnas problēmas ar standarta atbalstu [IFC], kas, iespējams, ir saistīts ar standarta datu modeļa loti lielo sarežģītību. Daļēji pie vainas šeit ir arī daži standarti, jo tajos dažas detaļas bieži vien nav definētas, pastāv liela brīvības pakāpe un dažādas iespējamās interpretācijas. Tie pieļauj lielu sarežģītību objektu organizācijā un glabāšanā, kas neveicina efektīvu universālu izpratni, unikālu ieviešanu un konsekventu datu modelēšanu [117].

- Atsauces pētījums par IFC programmatūras atbalstu, 2021. gads



#### 6.1-6. attēls Dažādi ģeometriskie kodoli sniedz dažādas tās pašas parametriski aprakstītās ģeometrijas reprezentācijas (pēc [117]).

Pareiza izpratne par "konkrētiem noteikumiem" ir pieejama to īpašo organizāciju dalībniekiem, kas izstrādā IFC. Līdz ar to tas, kurš vēlas piekļūt svarīgām zināšanām par noteiktām IFC iezīmēm, centīsies sadarboties ar lielajiem CAD- piegādātājiem vai panākt kvalitatīvu iezīmju apzināšanu, veicot savu izpēti

*Jums rodas jautājums par datu importēšanu un eksportēšanu, izmantojot IFC formātu, un jūs uzdodat to citiem piegādātājiem: "Kāpēc IFC failā ir informācija par telpu parametrisko pārnesi? Atvērtajā specifikācijā par to nekas nav teikts". Atbilde no "zinošākiem" Eiropas piegādātājiem: "Jā, tas nav pateikts, bet tas ir atļauts".*

- No intervijas ar CAD 2021 izstrādātāju [118].

IFC apraksta ģeometriju, izmantojot parametriskos primitīvus, bet nesatur iebūvētu kodolu - tā lomu pilda CAD programma, kas apkopo ģeometriju, izmantojot ģeometrijas kodolu. Ģeometrijas kodols veic matemātiskos aprēķinus un definē krustpunktus, bet IFC tikai sniedz datus tā interpretācijai. Ja IFC satur nepareizas virsmas, dažādas programmas ar dažādiem ģeometrijas kodoliem atkarībā no kodola var tās ignorēt vai radīt klūdas.

Rezultātā, lai strādātu ar IFC formātu, ir jāatbild uz galveno jautājumu, uz kuru ir grūti atrast viennozīmīgu atbildi - kāds rīks, ar kādu ģeometrisko kodolu ir jāizmanto, lai iegūtu tādu datu kvalitāti, kāda sākotnēji bija projektā CAD programmā, no kurās tika iegūts IFC?

Datu kvalitātes jautājumi un IFC formāta sarežģītība neļauj tieši izmantot projektēšanas datus procesu automatizācijai, analīzei un datu apstrādei, kas bieži novēd izstrādātājus pie neizbēgamas nepieciešamības izmantot slēgtus CAD -risinājumus ar "kvalitatīvu" piekļuvi datiem[63], par ko rakstīja pats ražotājs, kurš IFC reģistrēja 1994. gadā [65].

Visas IFC parametru kartēšanas un ģenerēšanas īpatnības ģeometrijas kodolā var realizēt tikai lielas izstrādātāju komandas, kurām ir pieredze darbā ar ģeometrijas kodoliem. Tāpēc pašreizējā IFC formāta īpatnību un sarežģītības prakse ir izdevīga galvenokārt CAD- pārdevējiem un tai ir daudz kopīga ar lielo programmatūras pārdevēju stratēģiju "adopt, extend, destroy", kad augošā standarta sarežģītība faktiski rada šķēršļus mazajiem tirgus dalībniekiem [94].

Lielo pārdevēju stratēģija var būt pielāgot atvērtos standartus, pievienot patentētus paplašinājumus un funkcijas, lai radītu lietotāju atkarību no saviem produktiem un pēc tam izspiestu konkurentus.

IFC formāts, kas paredzēts kā universāls tilts starp dažādām CAD- (BIM-) sistēmām, patiesībā pilda dažādu CAD platformu ģeometrisko kodolu savietojamības problēmu indikatora lomu, līdzīgi kā STEP formāts, no kura tas sākotnēji radās.

Rezultātā šobrīd pilnīgu un kvalitatīvu IFC ontoloģijas ieviešanu ir iespējams īstenot lielajiem CAD piegādātājiem, kuri var ieguldīt ievērojamus resursus, lai atbalstītu visas vienības un to kartēšanu savā iekšējā ģeometrijas kodolā, kas nepastāv IFC kā standartam. Lielajiem piegādātājiem ir arī iespēja savstarpejī saskaņot tehnisko informāciju par funkcijām, kas var nebūt pieejama pat visaktīvākajam IFC formāta izstrādes organizāciju dalībniekiem.

Mazām neatkarīgām komandām un atvērtā koda projektiem, kas cenšas atbalstīt sadarbspējīgu formātu izstrādi, iekšēja ģeometrijas kodola trūkums kļūst par nopietnu problēmu. Bez tā praktiski nav iespējams ļemt vērā visas dažādās nianes un smalkumus, kas saistīti ar datu apmaiņu starp platformām.

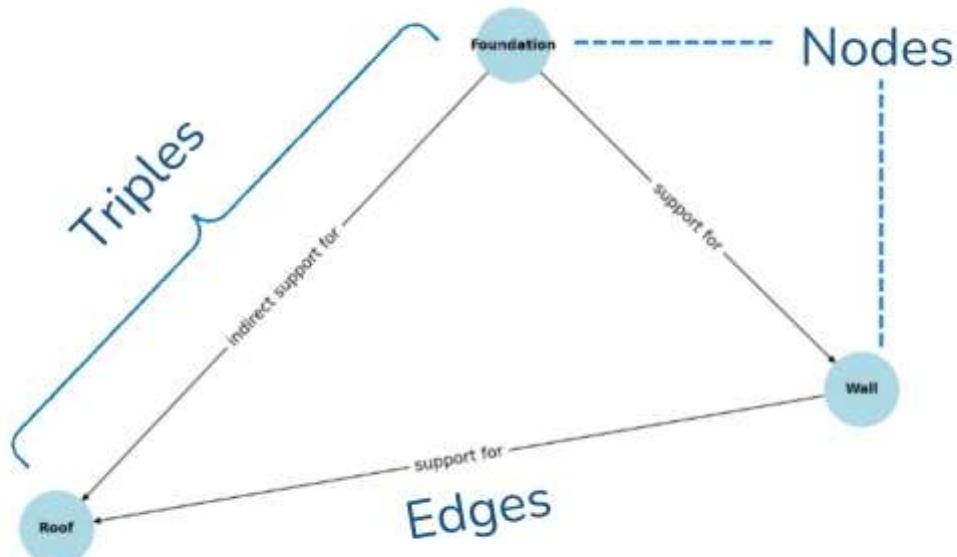
Līdz ar IFC parametriskā formāta un atvērtās BIM koncepcijas attīstību būvniecības nozarē ir saasinājušās diskusijas par ontoloģijas un semantikas lomu datu un procesu pārvaldībā.

## **Parādīšanās semantikas tēmas konstrukcijā un ontoloģijā**

Pateicoties idejām par semantisko internetu pagājušā gadsimta 90 gadu beigās un IFC formāta izstrādē iesaistīto organizāciju centieniem, semantika un ontoloģijas ir kļuvušas par vieniem no galvenajiem elementiem standartizācijā, par ko būvniecības nozarē diskutē līdz 20. gadsimta 20. gadu vidum.

Semantiskās tehnoloģijas ir lielu neviendabīgu datu masīvu unifikācija, standartizācija un modifīcēšana, kā arī sarežģītu meklējumu īstenošana.

OWL (Web Ontology Language), kas attēlota kā RDF grafi -triplets (Resource Description Framework) ( 6.1-7. attēls), tiek izmantota semantisko datu glabāšanai. OWL atsaucas uz grafu datu modeļiem, kuru veidus sīkāk aplūkojām nodaļā "Datu modeļi: datu attiecības un attiecības starp elementiem".



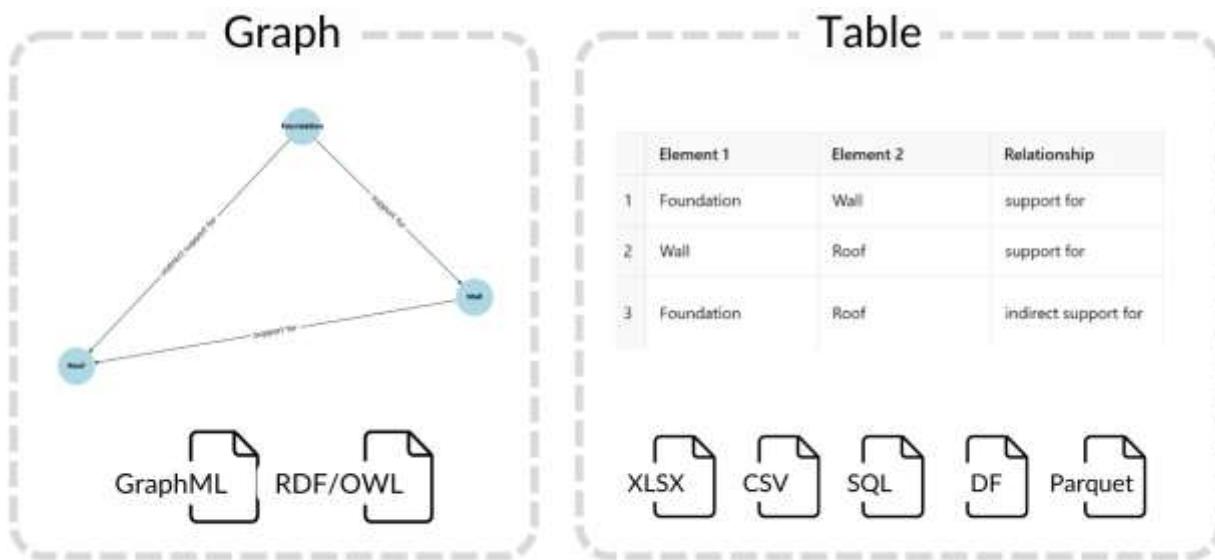
Attēls 6.1-7 RDF datu modelis: mezgli, malas un trīskārtas, kas ilustrē attiecības starp blokiem.

Teorētiski risoneru (automātiskās loģiskās secināšanas programmatūras) loģiskā secināšana ļauj no ontoloģijām iegūt jaunus apgalvojumus. Piemēram, ja būvniecības ontoloģijā ir ierakstīts, ka "pamats ir sienas balsts" un "siena ir jumta balsts" (6.1-7. attēls), risoneris spēj automātiski secināt, ka "pamats ir jumta balsts".

Šāds mehānisms ir noderīgs datu analīzes optimizēšanai, jo tas ļauj izvairīties no nepārprotamas visu atkarību noteikšanas. Tomēr tas nerada jaunas zināšanas, bet tikai identificē un strukturē jau zināmus faktus.

Semantika pati par sevi nerada jaunu nozīmi vai zināšanas, un šajā aspektā tā nav pārāka par citām datu glabāšanas un apstrādes tehnoloģijām. Relāciju datubāzu datu attēlošana kā tripli nepadara tos jēgpilnākus. Tabulu aizstāšana ar grafu struktūrām var būt noderīga datu modeļu unifikācijai, ērtai izguvei un drošai rediģēšanai, bet tā nepadara datus "gudrākus" - dators nesāk labāk saprast to saturu.

Datu loģiskās attiecības var organizēt bez sarežģītām semantiskām tehnoloģijām (6.1-8. attēls). Tradicionālās relāciju datu bāzes (SQL), kā arī CSV vai XLSX formāti ļauj veidot līdzīgas atkarības. Piemēram, kolonveida datubāzē var pievienot lauku "jumta atbalsts" un, veidojot sienu, automātiski saistīt jumtu ar pamatu. Šī pieeja tiek īstenota, neizmantojot RDF, OWL, grafikus vai risonerus, paliekot vienkāršs un efektīvs risinājums datu glabāšanai un analīzei.



Attēls 6.1-8 Grafa un tabulas datu modeļu salīdzinājums vienādu loģisko attiecību attēlošanai.

Vairāku lielo būvniecības uzņēmumu un IFC formāta izstrādes organizācijas [94] lēmums sekot semantiskā tīmekļa koncepcijai, kas 90. gadu beigās šķita daudzsološa, ir būtiski ietekmējis standartu izstrādi būvniecības nozarē.

Tomēr paradoxs ir tas, ka pati semantiskā tīmekļa koncepcija, kas sākotnēji bija paredzēta internetam, nav plaši izplatīta pat savā vietējā vidē. Neraugoties uz RDF un OWL attīstību, pilnvērtīgs semantiskais tīmeklis nav parādījies savā sākotnējā koncepcijā, un tā izveide jau ir maz ticama.

### Kāpēc semantiskās tehnoloģijas būvniecības nozarē neattaisno cerības?

Citas nozares ir saskārušās ar semantikas izmantošanas tehnoloģiju ierobežojumiem. Spēļu nozarē mēģinājumi aprakstīt spēļu objektus un to mijiedarbību, izmantojot ontoloģijas, izrādījās neefektīvi, jo to izmaiņu dinamika ir ļoti liela. Rezultātā priekšroka tika dota vienkāršākiem datu formātiem, piemēram, XML un JSON, kā arī algoritmiskiem risinājumiem. Līdzīga situācija bija nekustamā īpašuma nozarē: reģionālo terminoloģijas atšķirību un biežo tirgus izmaiņu dēļ ontoloģiju izmantošana izrādījās pārāk sarežģīta, savukārt vienkāršas datubāzes un tādi standarti kā RETS [119] labāk spēja tikt galā ar datu apmaiņas problēmām.

Tehniskās grūtības, piemēram, iezīmēšanas sarežģītība, darbietilpīgs atbalsts un zema izstrādātāju motivācija, palēnināja semantiskā tīmekļa ieviešanu un citās ekonomikas nozarēs. RDF (Resource Description Framework) nekļuva par masveida standartu, un ontoloģijas izrādījās pārāk sarežģītas un ekonomiski nepamatotas.

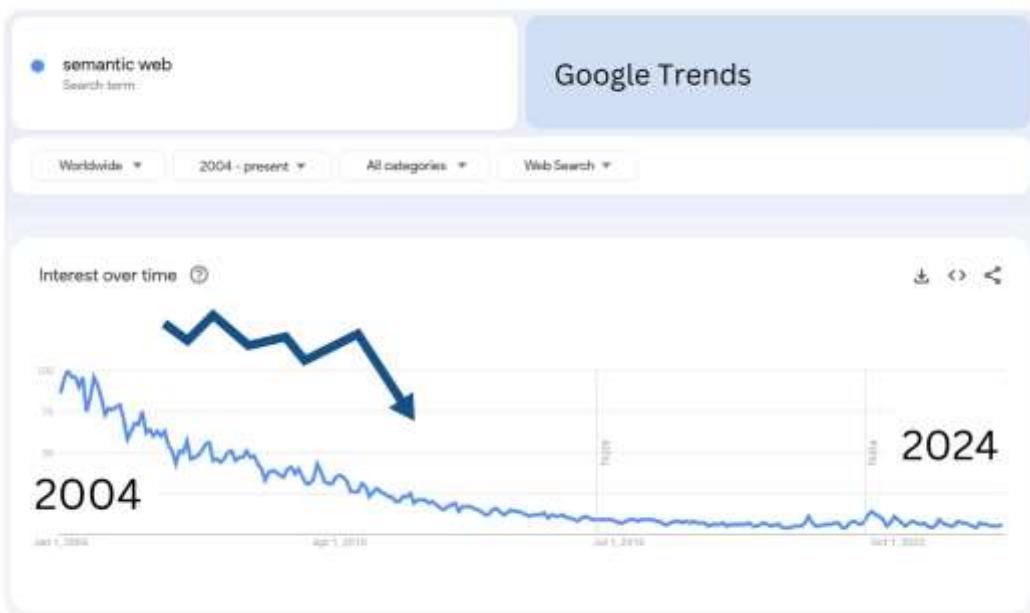
Tā rezultātā vērienīgā ideja par globāla semantiskā tīmekļa izveidi netika īstenota. Lai gan daži tehnoloģijas elementi, piemēram, ontoloģijas un SPARQL, ir nonākuši uzņēmumu risinājumos, sākotnējais mērķis izveidot vienotu visaptverošu datu struktūru nav sasniegts.

Koncepcija par internetu, kurā datori spēj saprast saturu, ir izrādījusies tehniski sarežģīta un komerciāli nerentabla. Tāpēc uzņēmumi, kas atbalstīja šo ideju, galu galā samazināja tās izmantošanu līdz atsevišķiem noderīgiem rīkiem, atstājot RDF un OWL ļoti specializētām uzņēmumu vajadzībām, nevis internetam kopumā. Pēdējo 20 gadu laikā veiktā Google tendenču analīze (6.1-9. attēls) liecina, ka semantiskajam tīmeklim, iespējams, vairs nav perspektīvu.

*Nav nepieciešams nevajadzīgi pavairot vienības. Ja ir vairāki loģiski konsekventi un vienlīdz labi izskaidrojoši kādas parādības skaidrojumi, tad, nemot vērā visus pārējos apstākļus, priekšroka jādod vienkāršākajam no tiem.*

žiletes princips

Te rodas loģisks jautājums: kāpēc vispār būvniecībā izmantot triplelus, riserus un SPARQL, ja datus var apstrādāt, izmantojot populārus strukturētus pieprasījumus (SQL, Pandas, Apache®)? Uzņēmumu lietojumprogrammās SQL ir standarts darbam ar datu bāzēm. SPARQL, gluži pretēji, prasa sarežģītas grafu struktūras un specializētu programmatūru, un, kā liecina Google tendencies, tā neiesaista izstrādātāju interesi.



Attēls 6.1-9 Interese par "semantiskā interneta" pieprasījumiem saskaņā ar Google statistiku.

Grafu datubāzes un klasifikācijas koki dažos gadījumos var būt noderīgi, taču to izmantošana ne vienmēr ir pamatota lielākajā daļā ikdienas uzdevumu. Rezultātā zināšanu grafiku izveide un semantiskā tīmekļa tehnoloģiju izmantošana ir lietderīga tikai tad, ja nepieciešams apvienot datus no dažādiem avotiem vai realizēt sarežģītus logiskus secinājumus.

Pāreja no tabulām uz grafiku datu modeļiem uzlabo meklēšanu un vienādo informācijas plūsmu, bet nepadara datus jēgpilnākus mašīnām. Jautājums nav par to, vai semantiskās tehnoloģijas būtu jāizmanto, bet gan par to, kur tās patiesām ir noderīgas. Pirms savā uzņēmumā ieviešiet ontoloģiju, semantiku un grafu datubāzes, noskaidrojet, kuri uzņēmumi jau veiksmīgi izmanto šīs tehnoloģijas un kur tās ir izgāzušās.

Neraugoties uz vērienīgajām cerībām, semantiskās tehnoloģijas tā arī nekļuva par universālu risinājumu datu strukturēšanai būvniecības nozarē. Praksē šīs tehnoloģijas nav radījušas universālu risinājumu, bet tikai pievienojušas jaunas sarežģītības, un šie centieni atkārto nerealizētos semantiskā interneta koncepcijas mērķus, kur cerības krietiņi pārsniedza realitāti.

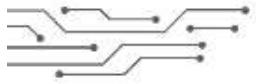


Attēls 6.1-10 Ģeometrija un informācija būvniecības procesos: no sarežģītām CAD un BIM-sistēmām līdz vienkāršotiem datiem analītikai.

Ja IT jomā semantiskā tīmekļa neveiksmes ir kompensētas ar jaunu tehnoloģiju parādīšanos (lielie dati, IoT, mašīnmācīšanās, AR/VR), tad būvniecības nozarē šādu iespēju nav.

Papildus izaicinājumiem, kas saistīti ar jēdzienu lietošanu, lai paziņotu par datu attiecībām starp projekta elementiem, joprojām pastāv būtiska problēma - šo datu pieejamība. Būvniecības nozarē joprojām dominē slēgtas sistēmas, kas apgrūtina darbu ar datiem, informācijas apmaiņu un procesu efektivitātes uzlabošanu.

Tieši datu slēgtība kļūst par vienu no galvenajiem šķēršļiem, kas kavē digitālo risinājumu attīstību būvniecībā. Atšķirībā no IT nozares, kur par standartu ir kļuvuši atvērti un saskaņoti datu formāti, CAD nozarē (BIM) katras programmatūra izmanto savu formātu, radot slēgtas ekosistēmas un mākslīgi ierobežojot lietotājus.



## NODAĻA 6.2.

### SLĒGTO PROJEKTU FORMĀTI UN SADARBSPĒJAS JAUTĀJUMI

#### Slēgti dati un samazināta produktivitāte: CAD nozares strupceļš (BIM)

CAD -sistēmu patentētais raksturs ir novēdis pie tā, ka katrai programmai ir sava unikāls datu formāts, kas ir vai nu slēgts un no ārpuses nepieejams - RVT, PLN, DWG, NDW, NWD, SKP, vai arī ir pieejams daļēji strukturētā formātā, izmantojot diezgan sarežģītu konvertēšanas procesu JSON, XML (CPIXML), IFC, STEP un ifcXML, IfcJSON, BIMJSON, IfcSQL, CSV utt..

Dažādie datu formāti, kuros var glabāt vienus un tos pašus datus par vieniem un tiem pašiem projektiem, atšķiras ne tikai pēc struktūras, bet ietver arī dažādas iekšējo iezīmju versijas, kas izstrādātājiem jāņem vērā, lai nodrošinātu lietojumprogrammu savietojamību. Piemēram, 2025. gada CAD formāts tiks atvērts 2026. gada CAD programmā, bet tas pats projekts nekad netiks atvērts visās CAD programmas versijās, kas, iespējams, bija pieejamas pirms 2025. gada.

Nenodrošinot tiešu piekļuvi datubāzēm, programmatūras piegādātājs būvniecības nozarē bieži vien izveido savu unikālu formātu un rīkus, kas profesionālim (projektēšanas inženierim vai datu pārvaldniekam) jāizmanto, lai piekļūtu datiem, importētu un eksportētu tos.

Rezultātā CAD pamata (BIM) un saistīto risinājumu (piemēram, ERP/PMIS)) pārdevēji pastāvīgi paaugstina produktu lietošanas cenas, un parastie lietotāji ir spiesti maksāt "komisijas maksu" katrā datu pārsūtīšanas posmā, izmantojot formātus [63]: par pieslēgšanu, importēšanu, eksportēšanu un darbu ar datiem, kurus lietotāji ir izveidojuši paši.

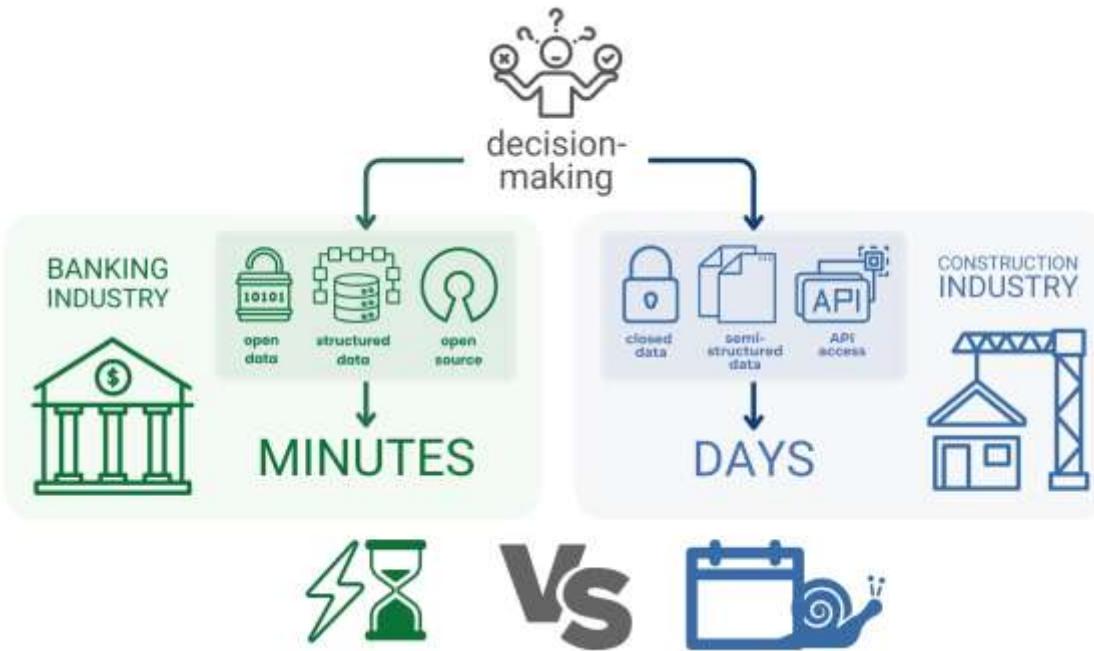
Populāru CAD - (BIM-) produktu datu piekļuves izmaksas mākoņkrātuvēs 2025. gadā sasniedgs 1 ASV dolāru par darījumu [120], bet būvniecības ERP -produktu abonēšanas maksa vidēja lieluma uzņēmumiem sasniedgs piecu un sešciparu summas gadā [121].

Mūsdienu būvniecības programmatūras būtība ir tāda, ka nevis automatizācija vai efektivitātes paaugstināšana, bet gan inženieru spēja izprast konkrētu augsti specializētu programmatūru ietekmē būvniecības projektu datu apstrādes kvalitāti un izmaksas, kā arī peļņu un uzņēmumu, kas īsteno būvniecības projektus, ilgtermiņa izdzīvošanu.

Piekļuves trūkums datu bāzēm CAD -sistēmām, kas tiek izmantotas desmitiem citu sistēmu un simtiem procesu [63], un no tā izrietošais kvalitatīvas komunikācijas trūkums starp atsevišķiem speciālistiem ir novēdis būvniecības nozari līdz vienas no neefektīvākajām ekonomikas nozarēm produktivitātes ziņā [44].

Pēdējo 20 gadu laikā, kopš tiek izmantotas CAD- (BIM-) projektēšanas lietojumprogrammas, parādās jaunas sistēmas (ERP), jaunas būvniecības tehnoloģijas un materiāli, visas būvniecības nozares produktivitāte ir samazinājusies par 20% (2.2.2-1. attēls), savukārt visu tautsaimniecības nozaru, kurās

nav būtisku problēmu ar piekļuvi datubāzēm un BIM koncepciju mārketingu, kopējā produktivitāte ir palielinājusies par 70% (apstrādes rūpniecībā - par 96%) [122].



Attēlā 6.2-1. Sakarā ar projektu datu izolētību un sarežģītību, no kuriem būvniecības nozarē ir atkarīgi desmitiem departamentu un simtiem procesu, lēmumu pieņemšanas ātrums ir vairākas reizes lēnāks nekā citās nozarēs.

Tomēr ir arī atsevišķi piemēri par alternatīvām pieejām sadarbspējas veidošanai starp CAD risinājumiem. Eiropas lielākais būvniecības uzņēmums ar SCOPE projektu [123], kas tika uzsākts jau 2018. gadā, demonstrē, kā ir iespējams pārsniegt klasisko CAD- (BIM-) sistēmu loģiku. Tā vietā, lai mēģinātu pakātot IFC vai paļautos uz patentētiem ģeometrijas kodoliem, SCOPE izstrādātāji izmanto API un SDK reversās inženierijas, lai iegūtu datus no dažādām CAD programmām, konvertētu tos neitrālos formātos, piemēram, OBJ vai CPIXML, pamatojoties uz vienīgo atvērtā koda ģeometrijas kodolu OCCT, un tālāk piemēro tos simtiem būvniecības un projektēšanas uzņēmumu uzņēmējdarbības procesu. Tomēr, neraugoties uz idejas progresivitāti, šādi projekti saskaras ar brīvo ģeometrisko kodolu ierobežojumiem un sarežģītību, un tie joprojām daļa no viena uzņēmuma slēgtām ekosistēmām, kas atkārto monopārdevēju risinājumu loģiku.

Sakarā ar slēgto sistēmu ierobežojumiem un datu formātu atšķirībām, kā arī efektīvu rīku trūkumu to unifikācijai, uzņēmumiem, kuriem jāstrādā ar CAD formātiem, nākas saskarties ar ievērojamu datu apjomu uzkrāšanu ar atšķirīgu struktūras un noslēgtības pakāpi. Šie dati netiek pienācīgi izmantoti un pazūd arhīvos, kur tie paliek mūžīgi aizmirsti un neizmantoti.

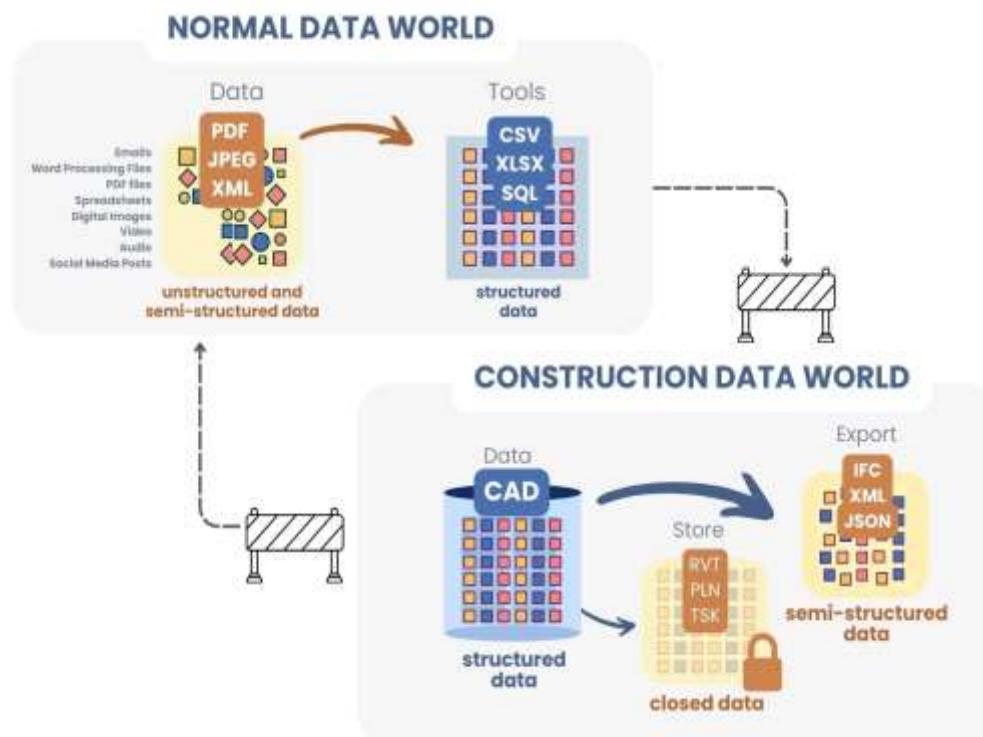
Dati, kas iegūti, ieguldot ievērojamas pūles projektēšanas posmā, kļūst nepieejami tālākai izmantošanai to sarežģītības un noslēgtā rakstura dēļ.

Tā rezultātā pēdējo 30 gadu laikā būvniecības nozares izstrādātāji ir bijuši spiesti atkal un atkal

saskarties ar vienu un to pašu problēmu: katrs jauns slēgta formāta vai patentēts risinājums rada nepieciešamību integrēties ar esošajām atvērtajām un slēgtajām CAD sistēmām. Šie nemitīgie mēģinājumi nodrošināt dažādu CAD un BIM risinājumu savietojamību tikai sarežģī datu ekosistēmu, nevis veicina tās vienkāršošanu un standartizāciju.

## Mīts par sadarbspēju starp CAD sistēmām

Ja deviņdesmito gadu vidū galvenais sadarbspējas attīstības virziens CAD vidē bija patentētā DWG formāta nojaukšana, kas vainagojās ar Open DWG alianses uzvaru [75] un populārākā rasējumu formāta faktisko atvēršanu visai būvniecības nozarei, tad līdz 20. gadsimta 20. gadu vidum uzmanības centrā ir mainījusies. Būvniecības nozarē arvien lielāku vērienu iegūst jauna tendence: daudzas izstrādes komandas koncentrējas uz tā saukto "tiltu" izveidi starp slēgtām CAD sistēmām (slēgtais BIM), IFC formātu un atvērtiem risinājumiem (atvērtais BIM). Lielākā daļa šo iniciatīvu balstās uz IFC formāta un OCCT ģeometrijas kodola izmantošanu, nodrošinot tehnisko tiltu starp atšķirīgām platformām. Šī pieeja tiek uzskatīta par daudzsološu virzienu, kas var būtiski uzlabot datu apmaiņu un programmatūras rīku sadarbspēju.

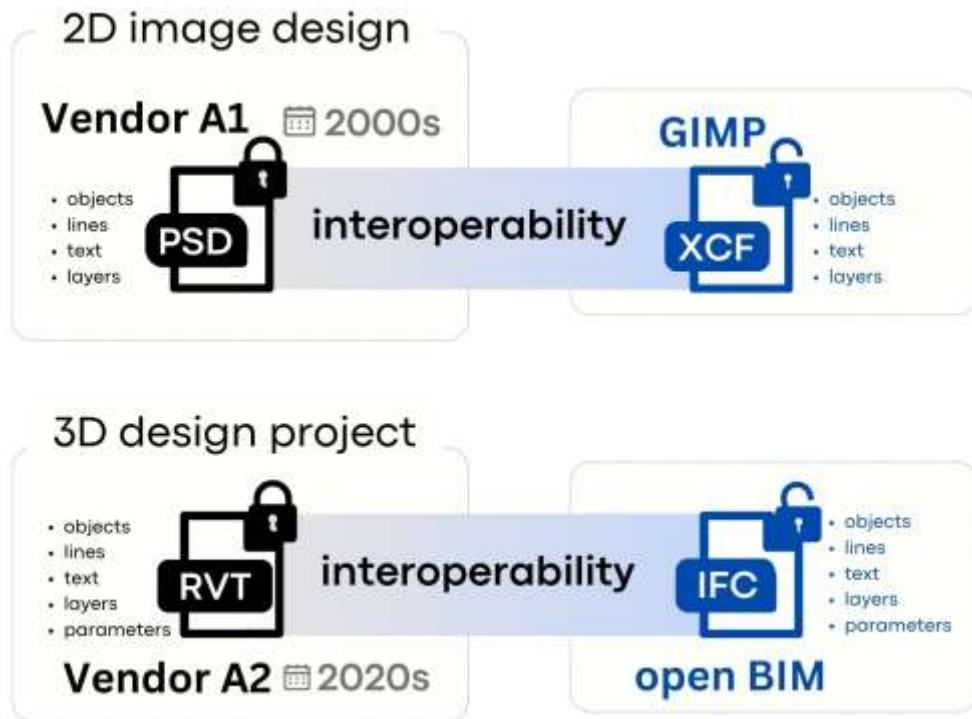


6.2-2. attēls Kamēr citas nozares strādā ar atvērtiem datiem, būvniecības nozarei nākas strādāt ar slēgtiem vai brīvi strukturētiem CAD formātiem (BIM).

Šai pieejai ir vēsturiskas paralēles. Divtūkstošajos gados izstrādātāji, cenšoties pārvarēt lielākā grafisko redaktoru ražotāja (2D world) dominējošo stāvokli, centās izveidot vienotu integrāciju starp savu patentēto risinājumu un bezmaksas atvērtā koda - alternatīvu GIMP (6.2-3. attēls). Toreiz, tāpat kā šodien būvniecībā, runa bija par mēģinājumu savienot slēgto un atvērto sistēmu, vienlaikus

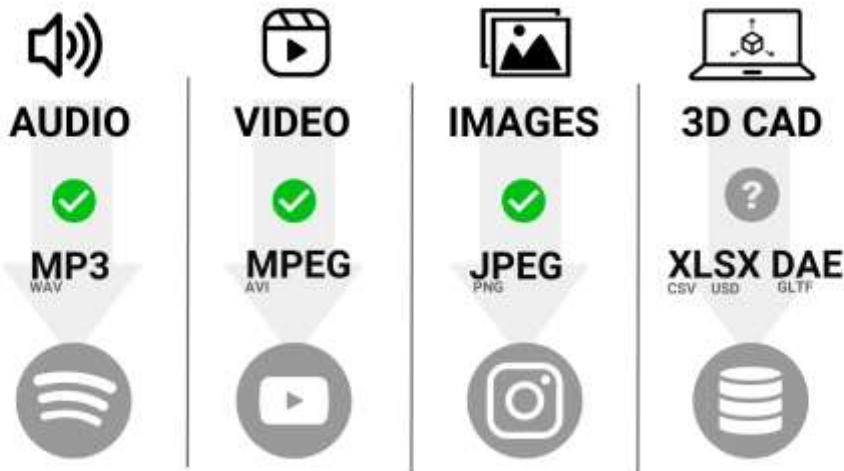
saglabājot sarežģītus parametrus, slāņus un iekšējo programmatūras loģiku.

Tomēr lietotāji patiesībā meklēja vienkāršus risinājumus - plakanus, atvērtus datus bez pārmērigi sarežģītiem slāņiem un programmas parametriem (analogi ģeometriskajam kodam CAD). Lietotāji meklēja vienkāršus un atvērtus datu formātus bez pārmērīgas loģikas. Par šādiem formātiem grafikā kļuva JPEG, PNG un GIF. Šodien tos izmanto sociālajos tīklos, tīmekļa vietnēs, lietojumprogrammās - tos ir viegli apstrādāt un interpretēt neatkarīgi no platformas vai programmatūras ražotāja.



Attēls 6.2-3 Datu formātu savietojamība būvniecībā ir līdzīga ceļam, kāds bija 2000. gados, kad tika mēģināts apvienot kāda populāra ražotāja patentētu produktu un atvērtā koda GIMP.

Rezultātā šodien gandrīz neviens attēlveidošanas nozarē neizmanto slēgtus formātus, piemēram, PSD vai atvērto XCF, lietojumprogrammās, sociālajos tīklos, piemēram, Facebook un Instagram, vai kā saturu tīmekļa vietnēs. Tā vietā lielākajā daļā uzdevumu tiek izmantoti plakanie un atvērtie JPEG, PNG un GIF formāti, lai tos būtu ērti lietot un plaši savietojami. Atvērtie formāti, piemēram, JPEG un PNG, ir kļuvuši par attēlu koplietošanas standartu, pateicoties to daudzpusībai un plašam atbalstam, kas atvieglo to lietošanu dažādās platformās. Līdzīga pāreja vērojama arī citos apmaiņas formātos, piemēram, video un audio formātos, kur uzsvars tiek likts uz tādiem universāliem formātiem kā MPEG un MP3 to kompresijas efektivitātes un plašās saderības dēļ. Šāda pāreja uz standartizāciju ir vienkāršojuši saturu un informācijas koplietošanu un atskanošanu, padarot tos pieejamus visiem lietotājiem dažādās platformās ( 6.2-4. attēls).



Attēls 6.2-4 Datu koplietošanai un izmantošanai ir kļuvuši populāri vienkāršoti formāti bez sarežģītām rediģēšanas funkcijām.

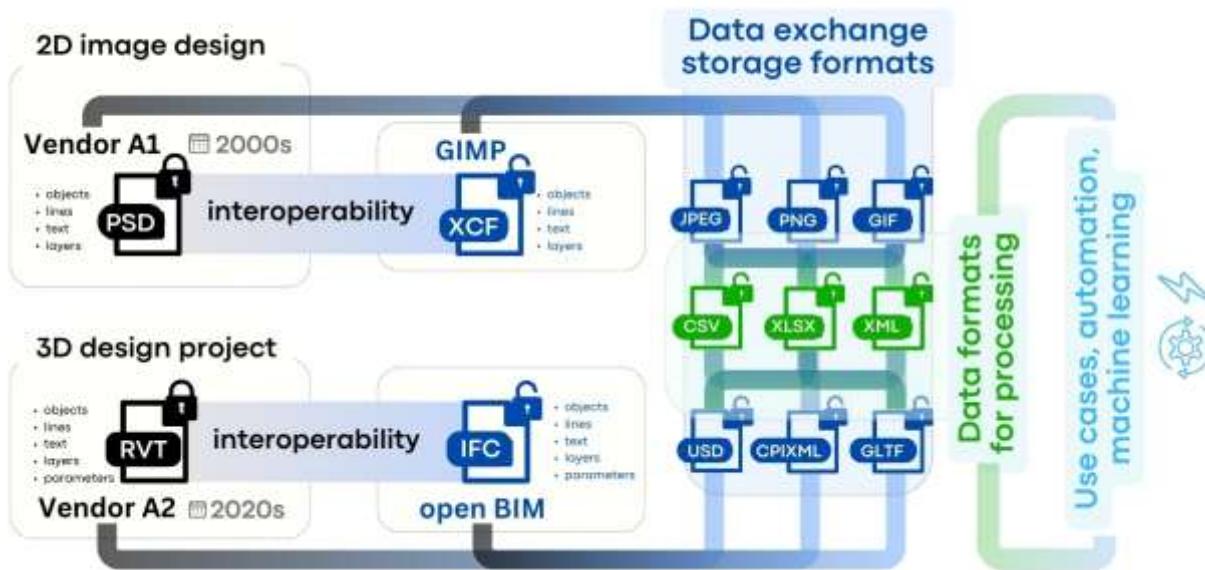
Līdzīgi procesi notiek arī 3D modelēšanā. Projektos datu apmaiņai ārpus CAD vides (BIM) aizvien biežāk tiek izmantoti tādi vienkārši un atvērti formāti kā USD, OBJ, glTF, DAE, DXF, SQL un XLSX. Šajos formātos tiek saglabāta visa nepieciešamā informācija, ieskaitot ģeometriju un metadatus, bez nepieciešamības izmantot sarežģītu BREP struktūru, ģeometrijas kodolus vai ražotāja specifiskus iekšējos klasifikatorus. Arī tādi patentēti formāti kā NWC, SVF, SVF2, CPIXML un CP2, ko nodrošina vadošie programmatūras ražotāji, veic līdzīgas funkcijas, taču atšķirībā no atvērtajiem standartiem tie ir slēgti.

Jāatzīmē (un ir vērts vēlreiz atgādināt, kā jau minēts iepriekšējā nodaļā), ka šo ideju - atteikšanos no starpposma neitrāliem un parametriskiem formātiem, piemēram, IGES, STEP un IFC - jau 2000. gadā atbalstīja lielākais CAD ražotājs, kas izveidoja BIM Whitepaper un 1994. gadā reģistrēja IFC formātu. Baltajā grāmatā 2000. gadā "Integrēta projektēšana un ražošana". [65] CAD ražotājs uzsver, cik svarīga ir dabiska piekļuve CAD datubāzei programmatūras vidē, neizmantojot starpposma tulkotājus un parametriskos formātus, lai saglabātu informācijas pilnīgumu un precizitāti.

Būvniecības nozarei vēl nav panākta vienošanās ne par rīkiem, ar kuriem piekļūt CAD datubāzēm vai to piespedu reverso inženieriju, ne par kopēja vienkāršota datu formāta pieņemšanu izmantošanai ārpus CAD platformām (BIM). Piemēram, daudzi lieli uzņēmumi Centrāleiropā un vācu valodā runājošos reģionos, kas darbojas būvniecības nozarē, izmanto CPIXML formātu savās ERP -sistēmās [121]. Šis patentētais formāts, kas ir sava veida XML, apvieno CAD (BIM) projekta datus, tostarp ģeometriskos un metadatus, vienā organizētā vienkāršotā struktūrā. Lieli būvniecības uzņēmumi arī paši veido jaunus formātus un sistēmas, piemēram, SCOPE projektā, par kuru mēs runājām iepriekšējā nodaļā.

Parametrisko CAD formātu vai sarežģīto parametrisko failu IFC (STEP) slēgtā loģika lielākajā daļā biznesa procesu ir lieka. Lietotāji meklē vienkāršotus un plakanus formātus, piemēram, USD, CPIXML,

XML &OBJ, DXF, glTF, SQLite, DAE &XLSX, kas satur visu nepieciešamo elementu informāciju, bet nav apgrūtināti ar lieku BREP ģeometrijas loģiku, atkarību no ģeometrijas kodoliem un konkrētu CAD un BIM -produktu iekšējām klasifikācijām (1. attēls. 6.2-5).



Attēls 6.2-5 Lielākajā daļā lietošanas gadījumu lietotāji izvēlas visvienkāršākos iespējamos formātus, kas nav atkarīgi no piegādātāja programmām.

Plašo attēlu formātu, piemēram, JPEG, PNG un GIF, atbrīvošana no pārpaliķušās ražotāju iekšējo dzinēju loģikas ir veicinājusi tūkstošiem sadarbspējīgu grafikas apstrādes un izmantošanas risinājumu izstrādi. Tā rezultātā radās dažādas lietojumprogrammas, sākot no retušas un filtrēšanas rīkiem un beidzot ar tādiem sociālajiem tīkliem kā Instagram, Snapchat un Canva, kur šos vienkāršotos datus var izmantot, nebūdam iepriekšējiem konkrētam programmatūras izstrādātājam.

Projektēšanas CAD -formātu standartizācija un vienkāršošana veicinās daudzu jaunu lietotājam draudzīgu un neatkarīgu rīku rašanos darbam ar būvniecības projektiem.

Pāreja no sarežģītās loģikas, ko izmanto piegādātāju lietojumprogrammas, kas piesaistītas slēgtiem ģeometrijas kodoliem, uz universāliem atvērtiem formātiem, kuru pamatā ir vienkāršotu elementu bibliotēkas, rada priekšnoteikumus elastīgākai, pārredzamākai un efektīvākai datu apstrādei. Tas arī paver piekļuvi informācijai visām būvniecības procesā iesaistītajām pusēm - no projektētājiem līdz pasūtītājiem un apkopes dienestiem.

Tomēr ļoti iespējams, ka tuvākajos gados CAD piedāvātāji mēģinās atkal novirzīt debates par sadarbspēju un piekļuvi CAD datubāzēm. Runa jau būs par "jaunām" koncepcijām - piemēram, granulāriem datiem, viediem grafikiem, "federatīviem modeļiem", digitālajiem dvīņiem mākoņu krātuvēs -, kā arī par nozares aliānsēm un standartiem, kas turpinās BIM un atvērtā BIM ceļu. Neraugoties uz pievilkīgo terminoloģiju, šādas iniciatīvas atkal var kļūt par rīkiem, lai noturētu lietotājus īpašumtiesību ekosistēmās. Viens no piemēriem ir USD (Universal Scene Description)

formāta aktīva popularizēšana kā "jaunā standarta" starpplatformu CAD (BIM) sadarbībai no 2023. gada.

## Pārejiet uz USD un granulāriem datiem m

AOUSD alianses izveidošana [124] 2023. gadā iezīmē nozīmīgu pavērsienu būvniecības nozarē. Mēs esam liecinieki tam, kā sākas jauna realitāte, ko veido CAD piedāvātāji, darbojoties ar būvniecības datiem, kas paredz vairākas būtiskas pārmaiņas. Pirmā lielā pārmaiņa attiecas uz CAD datu uztveri. Konceptuālā projektēšanas sākumposmā iesaistītie speciālisti arvien vairāk apzinās, ka projekta izveide CAD vidē ir tikai sākumpunkts. Projektēšanas procesā radītie dati galu galā kļūst par pamatu objektu analīzei, ekspluatācijai un pārvaldībai. Tas nozīmē, ka tiem jābūt pieejamiem un izmantojamiem sistēmās, kas nav tradicionālie CAD rīki.

Vienlaikus notiek apvērsums vadošo izstrādātāju pieejā. Nozares vadošais CAD- piegādātājs, BIM koncepčijas un IFC formāta radītājs, savā stratēģijā veic negaidītu pavērsienu. Sākot ar 2023. gadu, uzņēmums atsakās no tradicionālās datu uzglabāšanas atsevišķos failos, koncentrējoties uz darbu ar granulāriem (normalizētiem un strukturētiem) datiem un pārejot uz datu orientētu pieeju [125].

Pārdevēji seko citu nozaru vēsturiskajām tendencēm: lielākajai daļai lietotāju nav nepieciešami slēgti CAD formāti (līdzīgi PSD) vai sarežģīti parametriski IFC faili (līdzīgi GIMP ar slāņu loģiku). Viņiem ir nepieciešami vienkārši objektu attēli, kurus var izmantot CAFM (būvniecības Instagram), ERP (Facebook) un tūkstošiem citu procesu, kas piepildīti ar Excel izklājlapām un PDF dokumentiem.

Pašreizējās tendences būvniecības nozarē, iespējams, veido pamatu pakāpeniskai pārejai no parametriskiem un sarežģītiem formātiem uz universālākiem un neatkarīgākiem formātiem USD, GLTF, DAE, OBJ (ar metainformāciju gan hibrīdformātā, gan atsevišķos strukturētos vai brīvi strukturētos formātos). Vēsturiskie līderi, tostarp lielākie projektēšanas uzņēmumi, kas savulaik 90. gadu vidū aktīvi popularizēja IFC, tagad atklāti popularizē jauno USD formātu [93], uzsverot tā vienkāršību un daudzpusību (6.2-6. attēls). USD masveida ieviešana produktos, GLTF savietojamība un aktīva integrācija tādos rīkos kā Blender, Unreal Engine un Omniverse liecina par potenciāli jaunas paradigmas sākumu darbā ar datiem. Līdztekus tādu lokalizētu risinājumu popularitātei kā Eiropas plakanā USD formāts - CPIXML, ko izmanto populārājos Eiropas ERP, varētu potenciāli nostiprināt USD pozīcijas Centrāleiropā. IFC formāta izstrādē iesaistītās organizācijas jau pielāgo savu stratēģiju USD [126], kas tikai apliecina pārejas neizbēgamību.

Technical Specifications				Comparison / Notes
File Structure	Monolithic file	Uses ECS and linked data	IFC stores all data in one file; USD uses Entity-Component-System and linked data for modularity and flexibility	
Data Structure	Complex semantics, parametric geometry	Flat format, geometry in MESH, data in JSON	IFC is complex and parametric; USD is simpler and uses flat data	
Geometry	Parametric, dependent on BREP	Flat, MESH (triangular meshes)	IFC uses parametrics; USD uses meshes for simplified processing.	
Properties	Complex structure of semantic descriptions	Properties in JSON, easy access	Properties in USD are easier to use thanks to JSON	
Export/import	Complex implementation, dependent on third-party SDKs	Easy integration, wide support	USD integrates more easily and is supported in many products	
Format Complexity	High, requires deep understanding	Low, optimized for convenience	The time required to understand the structure of the file and the information stored in it.	
Performance	Can be slow when processing large models	High performance in visualization and processing	USD is optimized for speed and efficiency. Simulations, machine learning, AI, smart cities will be held in the Nvidia Omniverse	
Integration with 3D Engines	Limited	High, designed for graphics engines	USD excels with native support for real-time visualization platforms	
Support outside CAD Software	BlenderBIM, IfcOpenShell	Unreal Engine, Unity, Blender, Omniverse	USD is widely supported in graphics tools	
Cloud Technology Support	Limited	Well-suited for cloud services and online collaboration	USD is optimized for cloud solutions	
Ease of Integration into Web Applications	Difficult to integrate due to size and complexity	Easy to integrate, supports modern web technologies	USD is preferable for web applications	
Change Management	Versions through separate files	Versioning built into the format core	IFC handles changes via separate files, while USD embeds versioning directly into its structure	
Collaboration Support	Supports data exchange between project participants	Designed for collaborative work on complex scenes	USD provides efficient collaboration through layers and variations	
Learnability	Steep learning curve due to complexity	Easier to master thanks to a clear structure	USD is easier to learn and implement	

Attēls 6.2-6 IFC un USD formātu tehnisko specifikāciju salīdzinājums.

Nemot to vērā, USD ir potenciāls klūt par de facto standartu, kas sola pārvarēt daudzus pašreizējos ierobežojumus, kuri galvenokārt saistīti ar esošo CAD - (BIM-) formātu sarežģītību un to interpretācijas atkarību no ģeometriskiem kodoliem.

Parametrisko un sarežģīto CAD -formātu un IFC vietā būvniecības nozarē vienkāršo datu formātu USD, gLTF, DAE, OBJ ar elementu metainformāciju CSV, XLSX, JSON, XML vienkāršības un elastības dēļ ienems vietu vienkāršoti datu formāti USD, gLTF, DAE, OBJ ar elementu metainformāciju.

Pašreizējās pārmaiņas būvniecības nozarē pirmajā acu uzmetienā izskatās pēc tehnoloģiskā izrāviena, kas saistīts ar pāreju no novecojušās IFC uz modernāko USD. Tomēr ir vērts ņemt vērā, ka

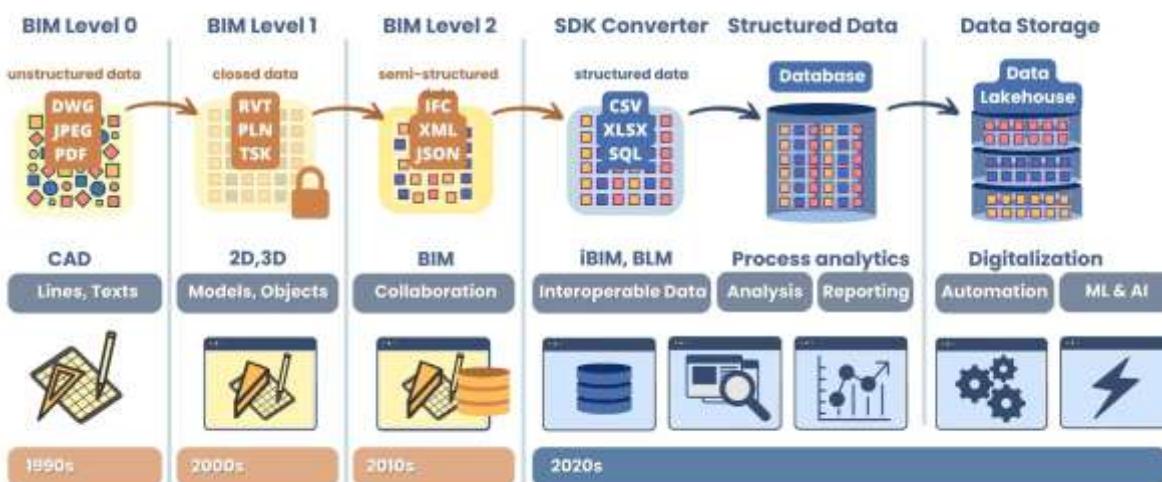
vēl 2000. gadā tas pats CAD ražotājs, kas izstrādāja IFC, rakstīja par tā problēmām un nepieciešamību pēc piekļuves datu bāzei [65], un tagad aktīvi veicina pāreju uz jaunu standartu - USD.

Aiz kārtējās "atvēto datu" fasādes USD un "jaunajām" koncepcijām granulārās datu pārvaldības jomā, izmantojot mākoņtehnoloģiju lietojumprogrammas, ko sāk popularizēt CAD pārdevēji, var slēpties pārdevēju nodoms monopolizēt projektu datu pārvaldību, kad lietotāji nonāk situācijā, kad formāta izvēle ir vairāk saistīta ar korporatīvajām interesēm nekā ar reālajām vajadzībām.

Galveno faktu analīze [93] rāda, ka šo izmaiņu galvenais mērķis drīzāk nav saistīts ar lietotāju ērtībām, bet galvenokārt ar ekosistēmu un datu plūsmu kontroles saglabāšanu par labu piegādātājiem, kuri 40 gadu laikā nekad nav spējuši nodrošināt piekļuvi CAD datubāzēm.

Iespējams, tagad ir pienācis laiks uzņēmumiem beigt gaidīt jaunas koncepcijas no programmatūras piegādātājiem un pievērsties pašattīstībai datu orientētā virzienā. Atbrīvojusies no datu piekļuves problēmām, izmantojot apgrieztās inženierijas rīkus, nozare varēs patstāvīgi pāriet uz mūsdienīgiem, bezmaksas un ērtiem rīkiem darbam un datu analīzei, neuzspiežot jaunas koncepcijas.

### CAD (BIM) Maturity Levels: From Stage 0 to Structured Data



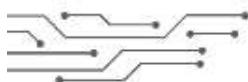
Attēls 6.2-7 CAD brieduma līmenis (BIM): no nestukturētiem datiem uz strukturētiem datiem un repozitorijiem.

Piekļuve datubāzēm, atvērtie dati un formāti būvniecības nozarē neizbēgami klūs par standartu, neatkarīgi no piegādātāju mēginājumiem šo procesu kavēt - tas ir tikai laika jautājums (6.2-7. attēls). Šīs pārejas tempos var ievērojami pieaugt, ja arvien vairāk speciālistu iepazīsies ar atvērtajiem formātiem, datubāzu rīkiem un pieejamajiem reversās inženierijas SDK, kas ļauj tieši piekļūt CAD datiem -sistēmām [92].

Nākotne ir atvērti, vienoti un analītiski pieejami dati. Lai izvairītos no atkarības no piegādātāju risinājumiem un nekļūtu par slēgtu ekosistēmu ķīlniekiem, būvniecības un inženiertehniskajiem uzņēmumiem agrāk vai vēlāk būs jāpaļaujas uz atvērtību un neatkarību, izvēloties formātus un risinājumus, kas nodrošina pilnīgu datu kontroli.

Dati, kas šodien tiek radīti būvniecības nozarē, nākotnē būs galvenais resurss, lai pieņemtu uzņēmējdarbības lēmumus. Tie kalpos kā stratēģiskā "degviela", kas veicinās būvniecības uzņēmumu attīstību un efektivitāti. Būvniecības nozares nākotne ir saistīta ar spēju strādāt ar datiem, nevis ar datu formātu vai modeļu izvēli.

Lai izprastu atšķirību starp atvērtajiem formātiem USD, glTF, DAE, OBJ un patentētiem parametriskiem CAD formātiem, ir svarīgi aplūkot vienu no sarežģītākajiem un būtiskākajiem datu elementiem vizualizācijā un projektēšanas aprēķinos - ģeometriju un tās ģenerēšanas procesus. Un, lai saprastu, kā ģeometriskie dati kļūst par pamatu analītikas un aprēķinu veikšanai būvniecībā, ir nepieciešams iedzījināties ģeometrijas ģenerēšanas, transformēšanas un uzglabāšanas mehānismos.



## NODAĻA 6.3.

### ĢEOMETRIJA BŪVΝIECĪBĀ: NO LĪNIJĀM LĪDZ KUBIKMETRIEM

#### Kad līnijas pārvēršas naudā jeb kāpēc celtniekiem nepieciešama ģeometrija

Ģeometrija būvniecībā ir ne tikai vizualizācija, bet arī pamats precīziem kvantitatīviem aprēķiniem. Projekta modelī ģeometrija papildina elementu parametru sarakstus ( 3.1-16. attēls) ar tādiem svarīgiem tilpuma parametriem kā garums, laukums un tilpums. Šo tilpuma parametru vērtības tiek aprēķinātas automātiski, izmantojot ģeometrijas kodolus, un tās ir sākumpunkts tāmēm, grafikiem un resursu modeļiem. Kā jau esam aplūkojuši šīs grāmatas 5. daļā un nodaļā "Izmaksu aprēķināšana un aplēšu veidošana būvprojektos", tieši objektu grupu tilpuma parametri no CAD -modeļiem ir mūsdienu ERP, PMIS -sistēmu pamatā/ Ģeometrijai ir būtiska nozīme ne tikai projektēšanas posmā, bet arī projekta īstenošanas vadībā, grafiku kontrolē, budžeta veidošanā un ekspluatācijā. Tāpat kā pirms tūkstošiem gadu, būvējot Ēģiptes piramīdas, projekta precīzitāte bija atkarīga no garuma mēriem, piemēram, elkoņiem un līkumiem, šodien ģeometrijas interpretācijas precīzitāte CAD -programmās tieši ietekmē rezultātu: sākot no budžeta un termiņiem līdz darbuzņēmēju izvēlei un piegādes loģistikai

Ļoti konkurētspējīgā un budžeta ierobežotā vidē tilpuma aprēķinu precīzitāte, kas ir tieši atkarīga no ģeometrijas, klūst par izdzīvošanas faktoru. Mūsdienu ERP -sistēmas ir tieši atkarīgas no pareiziem tilpuma parametriem, kas iegūti no CAD - un BIM -modeļiem. Tāpēc precīzs elementu ģeometriskais apraksts nav tikai vizualizācija, bet gan galvenais izmaksu un būvniecības laika pārvaldības instruments.

Vēsturiski ģeometrija ir bijusi galvenā inženierzinātņu saziņas valoda. Sākot ar līnijām uz papirusa un beidzot ar digitālajiem modeļiem, rasējumi un ģeometriskie attēli ir kalpojuši kā informācijas apmaiņas līdzeklis starp projektētājiem, meistariem un tāmētājiem. Pirms datoru parādīšanās aprēķinus veica manuāli, izmantojot lineālus un transportierus. Mūsdienās šis uzdevums ir automatizēts, pateicoties telpiskajai modeļēšanai: CAD programmatūras ģeometriskie kodoli pārvērš līnijas un punktus trīsdimensiju ķermenos, no kuriem automātiski tiek iegūti visi nepieciešamie parametri.

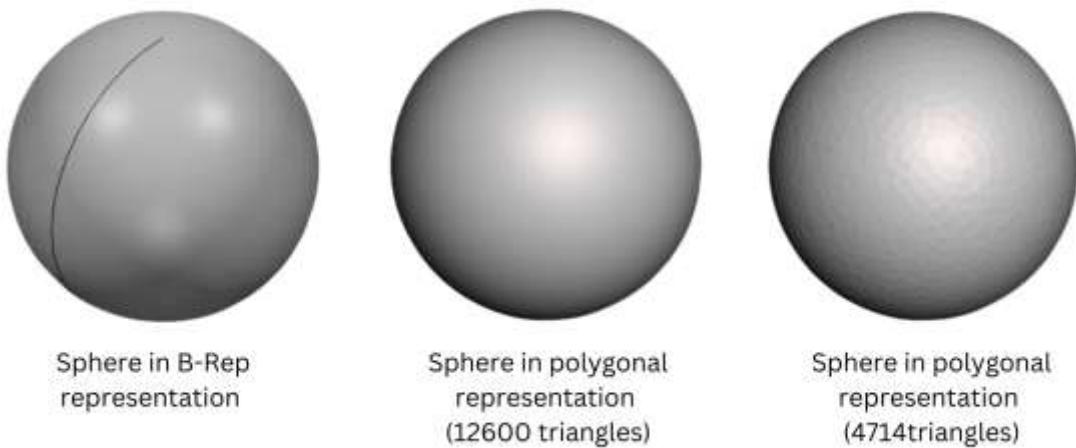
Strādājot CAD -programmās, ģeometrisko elementu izveide aprēķiniem tiek veikta, izmantojot CAD- (BIM-) programmu lietotāja saskarni. Lai pārveidotu punktus un līnijas par tilpuma ķermeniem, tiek izmantots ģeometriskais kodols, kas pilda galveno uzdevumu - ģeometrijas pārveidošanu tilpuma modeļos, no kuriem pēc aproksimācijas automātiski tiek aprēķināti elementa tilpuma raksturlielumi.

#### No līnijām uz apjomiem: kā platība un apjoms klūst par datiem

Inženiertehniskajā praksē apjomus un laukumus aprēķina no ģeometriskām virsmām, kas aprakstītas analītiski vai izmantojot parametriskos modeļus, piemēram, NURBS (nevienmērīgas racionālas B-splīnijas) BREP (robežu elementu attēlojuma) sistēmā.

NURBS (Non-Uniform Rational B-Splines) ir matemātisks veids, kā aprakstīt līknes un virsmas, savukārt BREP ir sistēma objekta pilnīgas trīsdimensiju ģeometrijas aprakstīšanai, ieskaitot tā robežas, ko var definēt, izmantojot NURBS.

Neraugoties uz BREP un NURBS precizitāti, tiem ir nepieciešami jaudīgi skaitļošanas resursi un sarežģīti algoritmi. Tomēr tiešie aprēķini no šādiem matemātiski precīziem aprakstiem bieži vien ir skaitliski sarežģīti, tāpēc praksē gandrīz vienmēr tiek izmantota teselācija - virsmu pārveidošana trīsstūru režīgi, kas vienkāršo turpmākos aprēķinus. Tesselation ir sarežģitas virsmas sadalīšana trīsstūros vai daudzstūros. CAD /CAE vidēs šo metodi izmanto vizualizācijai, apjoma aprēķiniem, sadursmju meklēšanai, eksportam uz tādiem formātiem kā MESH un sadursmju analīzei. Piemērs no dabas ir bišu bišu sātis, kur sarežģīta forma ir sadalīta regulārā režīgi ( 6.3-1. attēls).



**Attēls 6.3-1 Viena un tā pati sfēra parametriskajā aprakstā BREP un daudzstūra attēlojumā ar dažādu trīsstūru skaitu.**

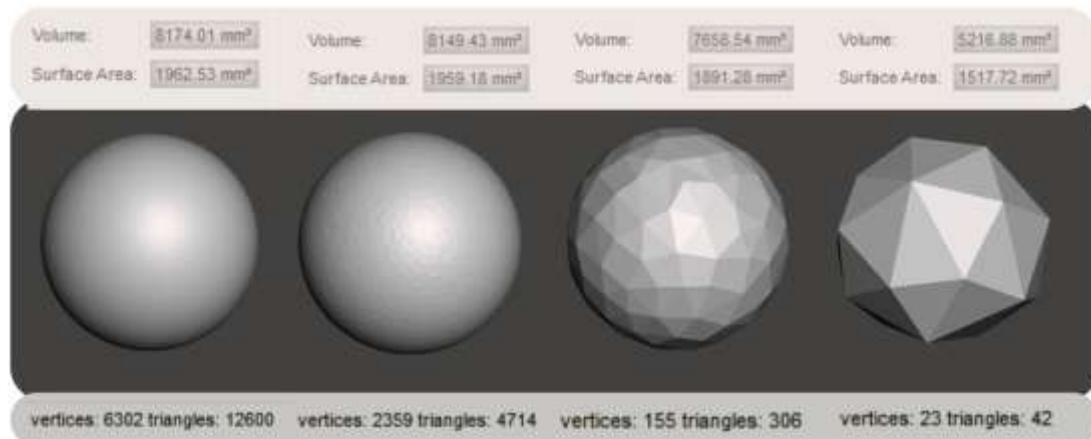
BREP (NURBS), ko izmanto CAD, nav ģeometrijas pamatmodelis. Tas tika izveidots kā ērts rīks apļu un racionālu splīniju attēlošanai un ģeometrijas datu uzglabāšanas samazināšanai. Tomēr tam ir ierobežojumi, piemēram, nespēja precīzi aprakstīt sinusoidu, kas ir pamatā spirālveida līnijām un virsmām, un nepieciešamība izmantot sarežģītus ģeometriskos kodolus.

Turpretī trīsstūrveida tīklus un parametrisku figūru teselēšanu raksturo vienkāršība, atmiņas efektīva izmantošana un spēja apstrādāt lielu datu apjomu ( 6.3-2. attēls). Šīs priekšrocības ļauj iztikt bez sarežģītiem un dārgiem ģeometriskiem kodoliem, kā arī desmitiem miljonu rindu koda, kas tajos iestrādāts, aprēķinot ģeometriskās formas.

Lielākajā daļā ēku gadījumu nav svarīgi, kā precīzi tiek definēti tilpuma raksturlielumi - ar parametisko modeļu (BREP, IFC) vai poligona (USD, glTF, DAE, OBJ) palīdzību. Ģeometrija paliek aptuvenā formā: vai nu izmantojot NURBS vai MESH, tas vienmēr ir aptuvens formas apraksts.

Ģeometrija, kas definēta kā daudzstūri vai BREP (NURBS), zināmā mērā ir tikai veids, kā aptuveni aprakstīt nepārtrauktu formu. Tāpat kā Freneļa integrāļiem nav precīzas analītiskas izteiksmes, arī ģeometrijas diskretizācija ar daudzstūru vai NURBS vienmēr ir aptuvens aprēķins, tāpat kā trīsstūrveida MESH.

Parametriskā ģeometrija BREP formātā ir nepieciešama galvenokārt gadījumos, kad ir svarīgs minimāls datu apjoms un kad tās apstrādei un attēlošanai var izmantot resursietilpīgus un dārgus ģeometrijas kodolus. Visbiežāk tas ir raksturīgi CAD -programmu izstrādātājiem, kas šim nolūkam savos produktos izmanto MCAD -piedāvātāju ģeometriskos kodolus. Šādā gadījumā pat šajās programmās BREP modeļi teselizācijas procesā vizualizācijai un aprēķiniem bieži tiek pārvērsti trīsstūros (līdzīgi kā PSD faili tiek vienkāršoti JPEG formātā).



**Attēls 6.3-2 Atšķirība starp tilpuma raksturlielumiem figūrās ar dažādu daudzstūru skaitu.**

Poligonālajam MESH, kā arī parametriskajam BREP, ir savas priekšrocības un ierobežojumi, taču mērķis ir viens - aprakstīt ģeometriju, nemot vērā lietotāja uzdevumu. Galu galā ģeometriskā modeļa precīzitāte ir atkarīga ne tikai no tā attēlojuma metodes, bet arī no konkrētā uzdevuma prasībām.

Lielākajā daļā konstruēšanas problēmu nepieciešamība pēc parametriskās ģeometrijas un sarežģītiem ģeometriskiem kodoliem var būt lieka.

Katrā konkrētā aprēķinu automatizācijas uzdevumā ir vērts apsvērt, vai parametriskās ģeometrijas nozīmi nav pārspīlējuši CAD izstrādātāji, kuri ir ieinteresēti savu programmatūras produktu popularizēšanā un pārdošanā.

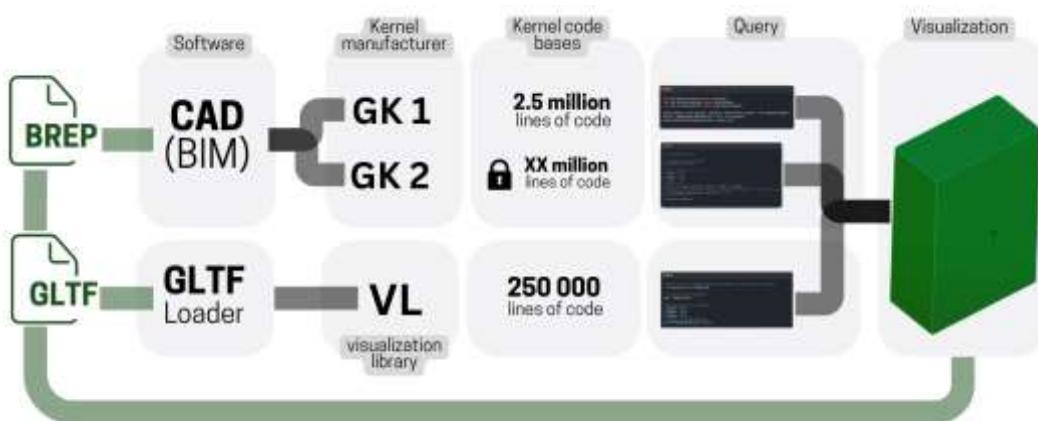
## Pāreja uz MESH, USD un daudzstūriem: ģeometrijai izmantojiet teselāciju

Būvniecības nozarē, straumējot, izstrādājot sistēmas, datubāzes vai automatizējot procesus, kas strādā ar projektēšanas informāciju un elementu ģeometriju, ir svarīgi censties panākt neatkarību no konkrētiem CAD redaktoriem un ģeometrijas kodoliem.

Apmaiņas formāts, kas izmantojams gan aprēķinu nodaļās, gan būvlaukumā, nedrīkst būt balstīts uz konkrētu CAD- (BIM-) programmu. Geometriskā informācija formātā būtu jāattēlo tieši, izmantojot teselēāciju, bez atsauces uz ģeometrisko kodolu vai CAD arhitektūru.

Parametrisko ģeometriju no CAD var uzskatīt par starpposma avotu, bet ne par universāla formāta pamatu. Lielākā daļa parametrisko aprakstu (tostarp BREP un NURBS) turpmākai apstrādei jebkurā gadījumā tiek pārvērsti daudzstūrveida MESH. Ja rezultāts ir vienāds (teselēācija un daudzstūri) un process ir vienkāršāks, izvēle ir acīmredzama. Tas ir analoģiski izvēlei starp grafu ontoloģijām un strukturētām tabulām (par ko mēs runājām ceturtajā daļā): pārmērīga sarežģītība reti kad ir pamatota (3.2-10. attēls, 6.1-8. attēls).

Atvērtajos formātos, piemēram, OBJ, STL, glTF, SVF, CPIXML, USD un DAE, tiek izmantota universāla trīsstūra acs struktūra, kas tiem dod būtiskas priekšrocības. Šiem formātiem ir lieliska savietojamība - tos ir viegli nolasīt un vizualizēt, izmantojot pieejamās atvērtā koda bibliotēkas, un tiem nav nepieciešami sarežģīti specializēti ģeometrijas kodoli, kas satur miljoniem rindu kodu (6.3-3. attēls). Šos daudzpusīgos ģeometrijas formātus izmanto dažādās lietojumprogrammās, sākot no salīdzinoši vienkāršiem virtuves projektēšanas rīkiem IKEA™ līdz sarežģītām objektu vizualizācijas sistēmām kino un VR -aplikācijās. Svarīga priekšrocība ir tas, ka darbam ar šiem formātiem ir pieejams liels skaits bezmaksas un atvērtā pirmkoda bibliotēku, kas ir pieejamas lielākajai daļai platformu un programmēšanas valodu.



Attēls 6.3-3 Tādu pašu ģeometrijas attēlojumu var panākt, izmantojot parametriskos formātus un ģeometrijas kodolus, vai izmantojot triangulētos formātus un atvērtā koda vizualizācijas bibliotēkas.

Tāpat kā paši lietotāji, arī CAD -pārdevēji saskaras ar problēmām, interpretējot ārzemju parametriskos CAD formātus vai atvērtos IFC dažādu ģeometrisko kodolu dēļ. Praksē visi bez izņēmuma CAD -pārdevēji datu pārsūtīšanai starp sistēmām izmanto reversās inženierijas SDK, un neviens no tiem savstarpējās savietojamības nolūkos nepajaujas uz tādiem formātiem kā IFC vai USD [93].

Tā vietā, lai izmantotu CAD- piegādātāju apvienību popularizētas koncepcijas, kurās tie paši neizmanto, CAD risinājumu izstrādātājiem un lietotājiem ir produktīvāk koncentrēties uz izpratni par katras pieejas priekšrocībām konkrētā kontekstā un izvēlēties vienu vai otru ģeometrijas veidu atkarībā no lietošanas gadījuma. Izvēle starp dažādiem ģeometriskiem attēlojumiem ir kompromiss starp precizitāti, skaitļošanas efektivitāti un konkrētā uzdevuma praktiskajām vajadzībām.

Sarežģītība, kas saistīta ar ģeometrisko kodolu izmantošanu, ko būvniecības nozarē tradicionāli uzspiež lielie piegādātāji, apstrādājot projektēšanas datus, bieži vien izrādās lieka. Uz MESH ģeometriju balstītais USD formāts var kļūt par sava veida "Pandoras lādi" nozarei, paverot izstrādātājiem jaunas iespējas organizēt datu apmaiņu - ārpus IFC un parametrisko BREP struktūru ietvara, kas raksturīgs CAD piegādātājiem.

Aplūkojot USD, DAE, gLTF, gLTF, OBJ u.c. formātu struktūru, kļūst skaidrs, ka pastāv vienkāršāki, atvērtāki formāti, kas ļauj efektīvi organizēt ģeometriskās informācijas pārsūtīšanu un izmantošanu bez nepieciešamības paļauties uz sarežģītiem parametriem un slēgtiem ģeometriskiem kodoliem. Šāda pieeja ne tikai pazemina tehnisko slieksni izstrādātājiem, bet arī veicina elastīgu, mērogojamu un patiesi atvērtu digitālās būvniecības risinājumu izstrādi.

### **LOD, LOI, LOMD - unikāla datu klasifikācija CAD (BIM).**

Pasaulē, kurā dažādās nozarēs tiek izmantoti dažādi datu detalizācijas un dzīluma līmeņi, papildus ģeometriskās attēlojuma formāti, CAD - (BIM-) metodoloģijas piedāvā savas unikālas klasifikācijas sistēmas, kas strukturē pieeju ēku modeļu informēšanai.

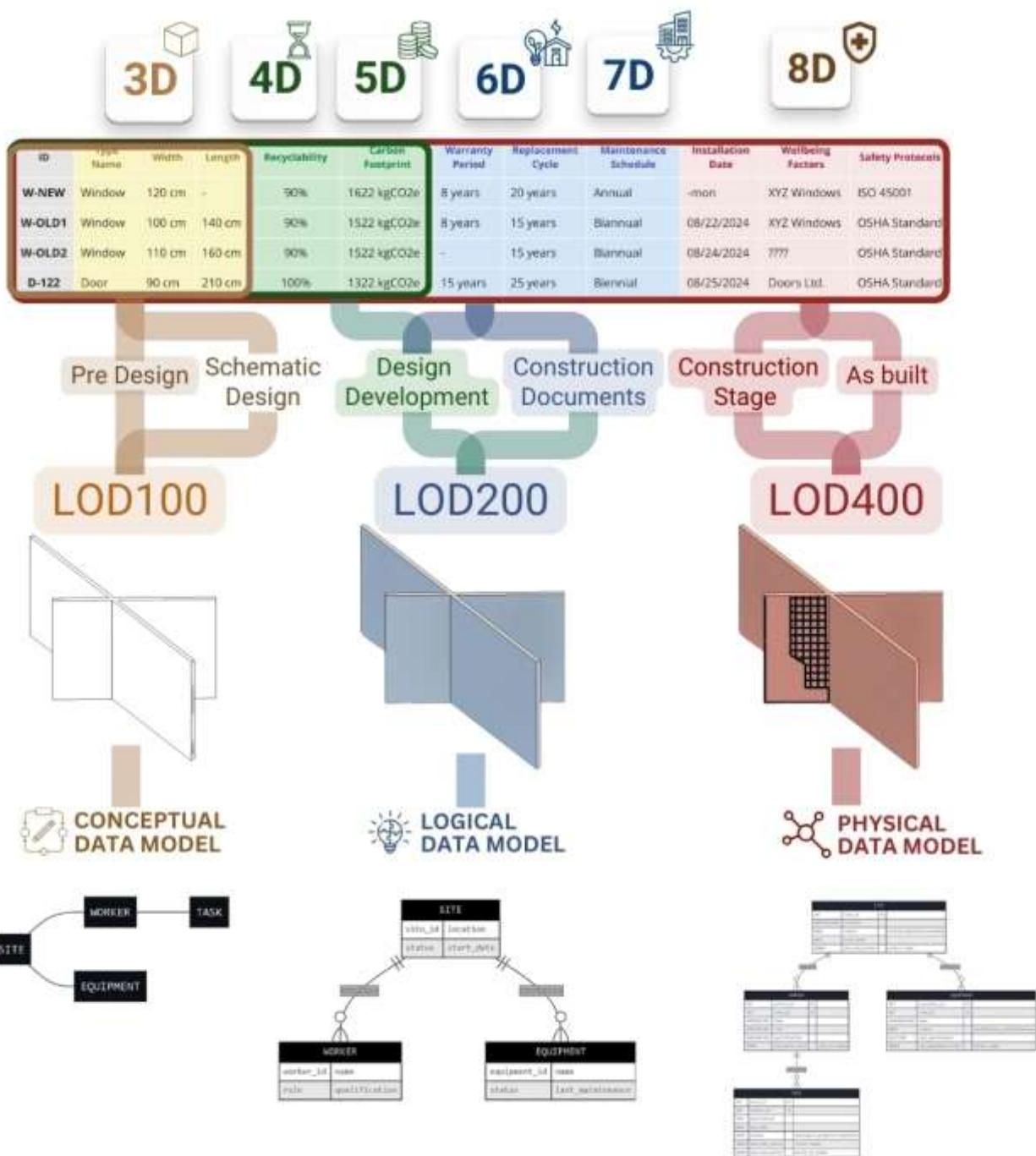
Viens no jaunās pieejas standartizācijai piemēriem ir modeļu izstrādes līmeņu ieviešana, kas atspoguļo gan grafisko, gan informācijas komponentu gatavības un uzticamības pakāpi. Informācijas saturā diferenciācijai darbā ar CAD - (BIM-) datiem parādījās LOD (Level Of Detail) - modeļa grafiskās daļas detalizācijas līmenis, un LOI (Level Of Information) - datu izstrādes līmenis. Turklāt integrētai pieejai tika ieviests jēdziens LOA (Level of Accuracy) - attēloto elementu precizitāte un LOG (Level of Geometry), lai noteiktu grafiskā attēlojuma precizitāti.

Detalizācijas līmeņi (LOD) ir norādīti ar skaitļiem no 100 līdz 500, kas atspoguļo modeļa izstrādes pakāpi. LOD 100 ir konceptuāls modelis ar vispārīgām formām un izmēriem. LOD 200 ietver precīzākus izmērus un formas, bet ar nosacītu detalizāciju. LOD 300 ir detalizēts modelis ar precīziem izmēriem, formām un elementu atrašanās vietām. LOD 400 ietver detalizētu informāciju, kas nepieciešama elementu izgatavošanai un uzstādīšanai. LOD 500 atspoguļo objekta faktisko stāvokli pēc būvniecības un tiek izmantots ekspluatācijai un apkopei. Šie līmeņi apraksta CAD (BIM) modeļa informācijas piesātinājuma struktūru dažādos dzīves cikla posmos, tostarp 3D, 4D, 5D un citos.

Reālos projektos augsta detalizācijas pakāpe (LOD400) bieži vien ir pārmērīga, un pietiek ar LOD100 ģeometriju vai pat plakaniem rasējumiem, bet pārējos datus var iegūt vai nu ar aprēķiniem, vai no saistītiem elementiem, kuriem var nebūt atšķirīgas ģeometrijas. Piemēram, telpām un telpu elementiem (telpu elementu kategorijas) var nebūt vizuālās ģeometrijas, bet tajās var būt ievērojams informācijas apjoms un datubāzes, ap kurām tiek veidoti daudzi uzņēmējdarbības procesi.

Tāpēc pirms projektēšanas uzsākšanas ir svarīgi skaidri definēt nepieciešamo detalizācijas pakāpi. 4D -7D izmantošanas gadījumiem bieži vien pietiek pat ar DWG rasējumiem un minimālu LOD100 ģeometriju. Galvenais uzdevums prasību procesā ir atrast līdzsvaru starp modeļa bagātību un praktiskumu.

Būtībā, ja CAD (BIM) datus uzskatām par datubāzi (kas tā arī ir), tad modeļa piesātinājuma apraksts ar jauniem akronīmiem nav nekas cits kā pakāpeniska informācijas sistēmu datu modelēšana, sākot no konceptuālā līmeņa un beidzot ar fizisko ( 6.3-4. attēls), kas tika detalizēti aplūkota grāmatas trešajā un ceturtajā daļā. Katrs LOD un LOI palielinājums nozīmē jaunas informācijas pievienošanu, kas nepieciešama jauniem uzdevumiem: aprēķiniem, būvniecības vadībai, ekspluatācijai, un to raksturo secīga modeļa bagātināšana ar papildu informācijas slāņiem (3D -8D) dažādu parametru veidā, ko mēs aplūkojām grāmatas piektajā daļā.



Attēls 6.3-4 Projekta detaļu precizēšanas process ir identisks datu modelēšanai no konceptuālā uz fizisko datu modeli.

Ģeometrija ir tikai daļa no projektēšanas datiem, kuru nepieciešamība būvprojektos ne vienmēr ir pamatota, un galvenais jautājums darbā ar CAD -datiem ir ne tik daudz tas, kā modeļi tiek vizualizēti, bet gan tas, kā šo modeļu datus var izmantot ārpus CAD (BIM) programmām.

Līdz 2000. gadu vidum būvniecības nozare saskārās ar vēl nebijušu izaicinājumu, jo strauji pieauga datu apjoms vadības un datu apstrādes sistēmās, jo īpaši to, kas nāk no CAD nodalām (BIM). Šis

krasais datu apjoma pieaugums pārsteidza uzņēmumu vadītājus, un viņi nebija gatavi pieaugošajām prasībām attiecībā uz datu kvalitāti un pārvaldību.

## **Jaunie CAD standarti (BIM) - AIA, BEP, IDS, LOD, COBie.**

Izmantojot to, ka trūkst brīvas piekļuves CAD datubāzēm un datu apstrādes tirgū ir ierobežota konkurence, kā arī izmantojot mārketinga kampaņas, kas saistītas ar jauno akronīmu BIM, organizācijas, kas izstrādā pieejas darbam ar CAD datiem, ir sākušas izstrādāt jaunus standartus un koncepcijas, kuru mērķis de jure ir uzlabot datu pārvaldības praksi.

Lai gan gandrīz visas CAD ražotāju un izstrādātāju tieši vai netieši atbalstītās iniciatīvas (BIM) ir bijušas vērstas uz darba plūsmu optimizāciju, to rezultātā ir izveidojusies virkne standartu, kurus lobē dažādas ieinteresētās puses, radot būvniecības nozarē zināmu neskaidrību un neskaidrību par datu procesiem.

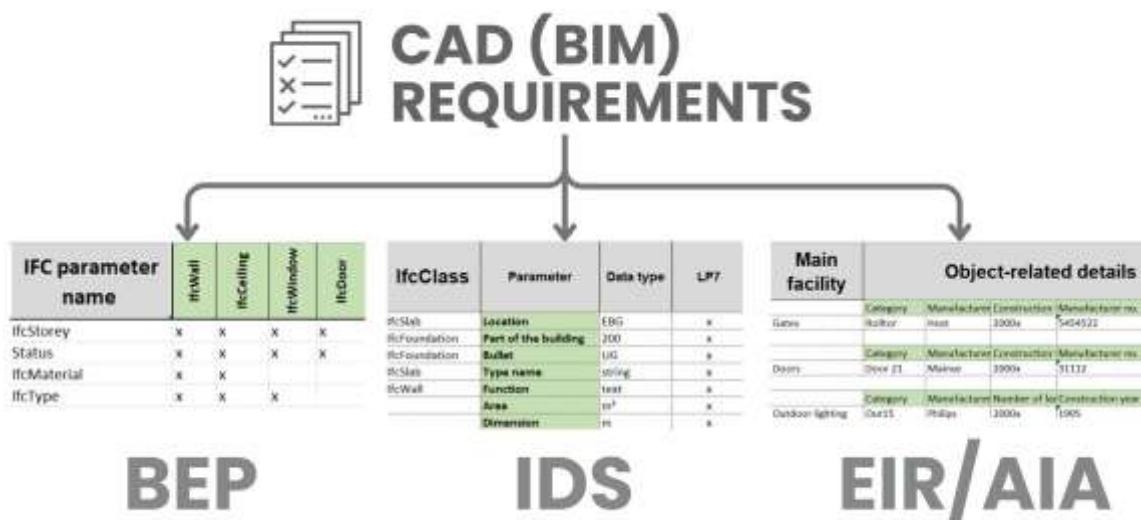
Uzskaitīsim dažus no jaunajiem datu standartiem papildus LOD, LOI, LOA, LOG, kas pēdējos gados parādījušies būvniecības nozarē:

- **BEP** (BIM Execution Plan) - apraksta, kā projektā integrēt un izmantot CAD (BIM), nosakot datu apstrādes metodes un procesus.
- **EIR dokuments /AIA** (Pasūtītāja prasības informācijai) - sagatavo pasūtītājs pirms konkursa izsludināšanas, un tajā ir iekļautas prasības, kas izpildītājam jāsagatavo un jāsniedz informācija. Tas kalpo par pamatu BEP attiecīgajā projektā.
- **AIM** (aktīvu informācijas modelis) ir daļa no BIM procesa. Kad projekts ir nodots un pabeigts, datu modeli sauc par aktīvu informācijas modeli jeb AIM. AIM mērķis ir pārvaldīt, uzturēt un ekspluatēt realizēto aktīvu.
- **IDS** (Informācijas piegādes specifikācija) - nosaka prasības un to, kādi dati un kādā formātā ir nepieciešami dažādos būvniecības projekta posmos.
- **iLOD** ir detalizācijas pakāpe LOD, ar kuru informācija ir attēlota BIM modelī. Tas nosaka, cik detalizēta un pilnīga ir modelī iekļautā informācija, sākot no pamata ģeometriskiem attēliem līdz detalizētām specifikācijām un datiem.
- **eLOD** - LOD atsevišķu elementu detalizācijas līmenis CAD modelī (BIM). Tas nosaka katras elementa modelēšanas pakāpi un ar to saistīto informāciju, piemēram, izmērus, materiālus, veikspējas raksturielumus un citus būtiskus atribūtus.
- **APS** (Platformas pakalpojumi) un citi lielāko CAD ražotāju produkti (BIM) - apraksta rīkus un infrastruktūru, kas nepieciešama saistītu un atvērtu datu modeļu izveidei.

Lai gan CAD (BIM) standartu, piemēram, LOD, LOI, LOA, LOG, BEP, EIR, AIA, AIM, IDS, iLOD, eLOD, ieviešanas deklarētais mērķis ir uzlabot datu pārvaldības kvalitāti un paplašināt automatizācijas iespējas, praksē to izmantošana bieži vien rada pārmērīgu sarežģītību un procesu sadrumstalotību. Ja CAD (BIM) modeli uzskatām par sava veida datu bāzi, kļūst skaidrs, ka daudzi no šiem standartiem

dublē citās nozarēs jau sen zināmas un efektīvas pieejas, ko izmanto darbā ar informācijas sistēmām. Tā vietā, lai vienkāršotu un unificētu, šādas iniciatīvas bieži vien rada papildu terminoloģisko slogu un kavē patiesi atvērtu un elastīgu risinājumu ieviešanu.

Jāatzīmē, ka daudzas no šīm jaunajām koncepcijām faktiski aizstāj modelēšanas un datu apstiprināšanas procesus, kas tika detalizēti aplūkoti grāmatas pirmajās daļās un kas jau sen tiek izmantoti citās ekonomikas nozarēs. Savukārt būvniecībā standartizācijas process bieži vien notiek pretējā virzienā - tiek radīti jauni datu aprakstīšanas formāti, jauni standarti un jaunas datu validēšanas koncepcijas, kas ne vienmēr nodrošina reālu vienotību un praktisku pielietojamību. Rezultātā tā vietā, lai vienkāršotu un automatizētu apstrādi, nozare saskaras ar papildu regulējuma un birokrātijas līmeņiem (6.3-1. attēls), kas ne vienmēr veicina efektivitātes paaugstināšanos.



Attēls 6.3-1 Datu un informācijas saturs prasības ir reducētas līdz atribūtu un to robežvērtību aprakstam, izmantojot tabulas.

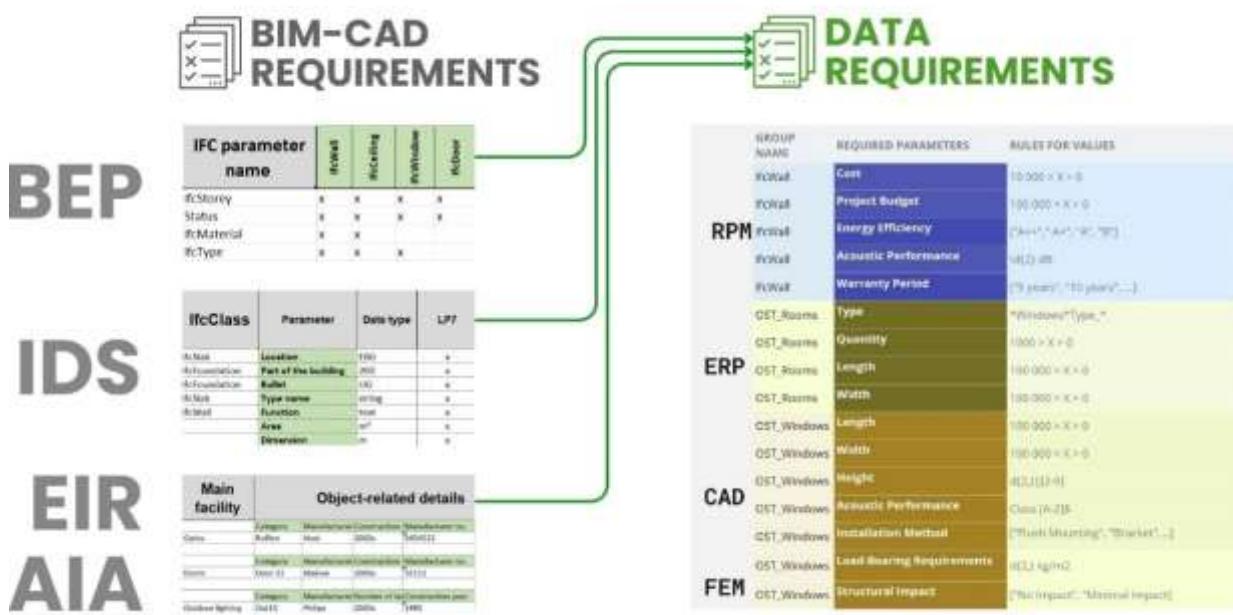
Tā vietā, lai vienkāršotu datu apstrādi, jaunie jēdzieni, kas saistīti ar CAD (BIM) datiem, visbiežāk rada papildu sarežģījumus un strīdus jau interpretācijas un pamatdefinīciju posmā.

Viens no jaunākajiem jauno koncepciju piemēriem ir IDS formāts (ieviests 2020. gadā), kas ļauj aprakstīt prasības informācijas modeļa atribūtu sastāvam atvērtajā BIM koncepcijā. IDS prasības apraksta informāciju par atribūtiem un to robežvērtībām strukturētas tabulas veidā (Excel vai MySQL), kas pēc tam tiek pārtulkota daļēji strukturēta XML formāta iezīmējumos, kas no XML pārdēvēts par īpašu saīsinājumu IDS.

Pretēji piegādātāju popularizētajam un BIM un atvērtā BIM atbalstītajam viedoklim, ka datu apstrāde būvniecībā ir unikāla, jo tiek izmantoti tādi specializēti rīki kā CAD un BIM, datu formāti un datu pārvaldības prakse šajā nozarē neatšķiras no citu nozaru datu apstrādes prakses

Prasību skaitu CAD projektiem un formātiem (BIM) var vienkāršot, izmantojot vienu prasību tabulu ar atribūtu kolonnām, kas detalizēti aprakstīta nodaļā "Prasību tulkošana strukturētā formā", un sākotnēji strukturētās prasības nav jātulko formātos, kas nav tabulas formāts (IDS sākotnēji tiek aprakstīta, izmantojot tabulu).

Vienkāršotā pieeja (6.3-2. attēls), kas ietver ailes entīžu identifikatoriem, īpašībām un robežvērtībām, kuras tika detalizēti aplūkotas iepriekšējās nodaļās (4.4-9. attēls, 4.4-16. attēls, 7.3-10. attēls), novērš nepieciešamību konvertēt prasības IDS-XML formātā. Šī metode nodrošina tiešu, mazāk apgrūtinošu un pārredzamāku datu kvalitātes kontroles mehānismu. Tā balstās uz plaši izmantotiem rīkiem, sākot no regulārām izteiksmēm (RegEx) līdz datu rāmjiem, Pandas un standarta ETL -payplanes - tieši tādiem pašiem, kādus izmanto citu tautsaimniecības nozaru speciālisti darbā ar datiem.



Attēls 6.3-2 Datu prasības citās nozarēs ir vienkāršotas līdz strukturētam atribūtu un to robežvērtību aprakstam.

Laika gaitā būvniecības nozarē datu slēgtā rakstura dēļ parādās arvien vairāk jaunu pieejumu un metožu, lai kontrolētu un pārvaldītu šos daudzveidīgos datus, lai gan dati būvniecības projektos būtībā ir tādi paši kā citās jomās. Kamēr citas nozares ir veiksmīgi iztikušas ar standartizētām datu apstrādes pieejām, būvniecībā turpina attīstīties jauni un unikāli datu formāti, prasības un apstiprināšanas koncepcijas.

Metodēm un instrumentiem, ko izmanto datu vākšanai, sagatavošanai un analīzei būvniecībā, nevajadzētu būt būtiski atšķirīgiem no tiem, ko izmanto citu tautsaimniecības nozaru speciālisti.

Nozarē ir izveidojusies atšķirīga terminoloģiskā ekosistēma, kas prasa kritisku pārdomu un pārvērtēšanu:

- STEP formāts ir pozicionēts ar jauno nosaukumu IFC, kas papildināts ar konstrukciju kategorizāciju, neņemot vērā paša STEP formāta ierobežojumus.
- Parametriskais formāts IFC tiek izmantots datu saziņas procesos, neraugoties uz to, ka trūkst vienota ģeometriskā kodola, kas nepieciešams vizualizācijai un aprēķiniem.
- Piekļuve datubāzēm CAD -sistēmām tiek popularizēta ar terminu "BIM", neapsriežot šo datubāzu specifiku un piekļuvi tām.
- Pārdevēji popularizē savietojamību, izmantojot IFC un USD formātus, bieži vien tos neīsteno praksē, izmantojot dārgu reverso inženieriju, ar ko paši ir saskārušies ar grūtībām.
- Termini LOD, LOI, LOA, LOG, BEP, EIR, AIA, AIM, IDS, eLOD tiek izmantoti universāli, lai aprakstītu vienus un tos pašus vienības parametrus, neatsaucoties uz modelēšanas un verifikācijas rīkiem, ko jau sen izmanto citās nozarēs.

Būvniecības nozare pierāda, ka būvniecības nozarē viss iepriekš minētais, lai gan tas izklausās dīvaini, ir iespējams, jo īpaši, ja galvenais mērķis ir gūt peļņu no katras datu apstrādes posma, pārdodot specializētus pakalpojumus un programmatūru. No uzņēmējdarbības viedokļa tajā nav nekā slikta. Tomēr paliek atklāts jautājums, vai šādi akronīmi un ar CAD (BIM) saistītas pieejas patiešām rada pievienoto vērtību un vienkāršo profesionālos procesus.

Būvniecības nozarē šāda sistēma darbojas, jo pati nozare šajā sistēmu un akronīmu labirintā gūst lielāko daļu savas spekulatīvās peļņas. Uzņēmumi, kas ir ieinteresēti pārredzamos procesos un atvērtos datos, ir reti sastopami. Šī sarežģītā situācija, visticamāk, turpināsies bezgalīgi - līdz klienti, klienti, investori, bankas un privātā kapitāla uzņēmumi sāks pieprasīt skaidrāku un informētāku pieejumu informācijas pārvaldībai.

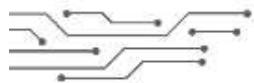
Nozarē ir uzkrāts pārmērīgs akronīmu skaits, taču tie visi dažādā mērā apraksta vienus un tos pašus procesus un datu prasības. To patiesā lietderība darba plūsmu vienkāršošanā joprojām ir apšaubāma.

Lai gan jēdzieni un mārketinga akronīmi nāk un zūd, datu prasību apstiprināšanas procesi uz visiem laikiem paliks neatņemama biznesa procesu sastāvdaļa. Tā vietā, lai radītu arvien vairāk specializētus formātus un noteikumus, būvniecības nozarei vajadzētu pievērsties rīkiem, kas jau ir pierādījuši savu efektivitāti citās jomās, piemēram, finanšu, rūpniecības un IT.

Terminu, akronīmu un formātu pārpilnība rada ilūziju par dziļi izstrādātiem digitālās būvniecības procesiem. Tomēr mārketinga jēdzieni un sarežģītā terminoloģija bieži slēpj vienkāršu, bet neērtu patiesību: dati joprojām ir grūti pieejami, slīkti dokumentēti un stingri piesaistīti konkrētiem programmatūras risinājumiem.

Lai izklītu no šī apburtā akronīmu un formātu formātu dēļ apburtā loka, ir nepieciešams uz CAD (BIM) sistēmām raudzīties nevis kā uz maģiskiem informācijas pārvaldības rīkiem, bet gan kā uz to, kas tās

patiesībā ir - specializētām datubāzēm. Tieši caur šo prizmu var saprast, kur beidzas mārketinga un sākas reāls darbs ar informāciju.



## NODAĻA 6.4.

### PROJEKTĒŠANAS PARAMETRU NOTEIKŠANA UN LLM IZMANTOŠANA KAD DARBĪBAI

#### CAD datu unikalitātes ilūzija (BIM): ceļš uz analītiku un atvērtiem formātiem

Modernās CAD (BIM) platformas ir būtiski pārveidojušas pieeju projektēšanas un būvniecības informācijas pārvaldībai. Ja agrāk šos rīkus galvenokārt izmantoja rasējumu un 3D modeļu izveidei, tad šodien tie darbojas kā pilnvērtīgas projekta datu krātuves. Saskaņā ar vienotā patiesības avota koncepciju parametriskais modelis arvien vairāk kļūst par galveno un bieži vienīgo projekta informācijas avotu, nodrošinot tā integrītāti un atbilstību visā projekta dzīves ciklā.

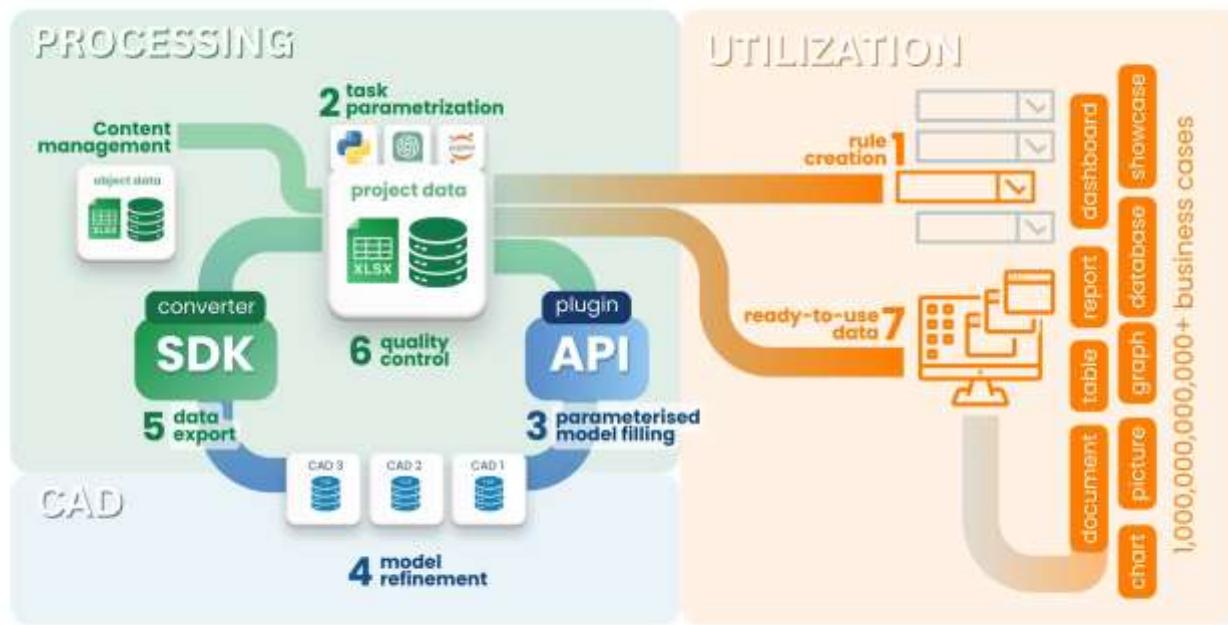
Galvenā atšķirība starp CAD - (BIM-) platformām un citām būvniecības datu pārvaldības sistēmām ir nepieciešamība pēc specializētiem rīkiem un API, lai piekļūtu informācijai (vienīgais patiesības avots). Šīs datubāzes nav universālas tradicionālajā izpratnē: atvērtās struktūras un elastīgas integrācijas vietā tās ir slēgta vide, stingri piesaistīta konkrētai platformai un formātam.

Neraugoties uz to, ka darbs ar CAD -datiem ir sarežģīts, ir svarīgāks jautājums, kas sniedzas tālāk par tehnisko realizāciju: kas īsti ir CAD datu bāzes (BIM)? Lai atbildētu uz šo jautājumu, ir jāiet tālāk par parastajiem akronīmiem un programmatūras izstrādātāju uzspiestajiem jēdzieniem. Tā vietā ir vērts pievērsties darba ar projekta informāciju būtībai - datiem un to apstrādei.

Biznesa process būvniecībā sākas nevis ar darbu ar CAD vai BIM rīkiem, bet gan ar projekta prasību veidošanu un datu modelēšanu. Vispirms tiek definēti uzdevuma parametri: vienību saraksts, to sākotnējie raksturlielumi un robežvērtības, kas jāņem vērā, risinot konkrētu uzdevumu. Tikai pēc tam CAD (BIM) sistēmās tiek izveidoti modeļi un elementi, pamatojoties uz noteiktajiem parametriem

Process, kas notiek pirms informācijas izveides CAD - (BIM)- datubāzēs, ir pilnīgi tāds pats kā datu modelēšanas process, kas detalizēti tika aplūkots grāmatas ceturtajā daļā un nodaļā "Datu modelēšana: konceptuālais, loģiskais un fiziskais modelis" ( 4.3-1. attēls).

Tāpat kā datu modelēšanā mēs veidojam prasības datiem, kurus vēlāk vēlamies apstrādāt datubāzē, CAD datubāzā vadītāji izveido projektēšanas prasības vairāku tabulu kolonnu vai atslēgu-vērtību pāru sarakstu veidā ( 6.4-1. attēls, 1.-2. solis). Un tikai pamatojoties uz šiem sākotnējiem parametriem, izmantojot API automātiski vai manuāli, projektētājs izveido (vai drīzāk precizē) objektus CAD- (BIM) datubāzēs (3.-4. solis), pēc tam vēlreiz pārbauda to atbilstību sākotnējām prasībām (5.-6. solis). Šis process - definēšana→ izveide→ validēšana→ pielāgošana (2.-6. posms) - atkārtojas iteratīvi, līdz datu kvalitāte, tāpat datu modelēšanā, sasniedz mērķa sistēmai - dokumentiem, tabulām vai paneļiem - vēlamo līmeni (7. posms).

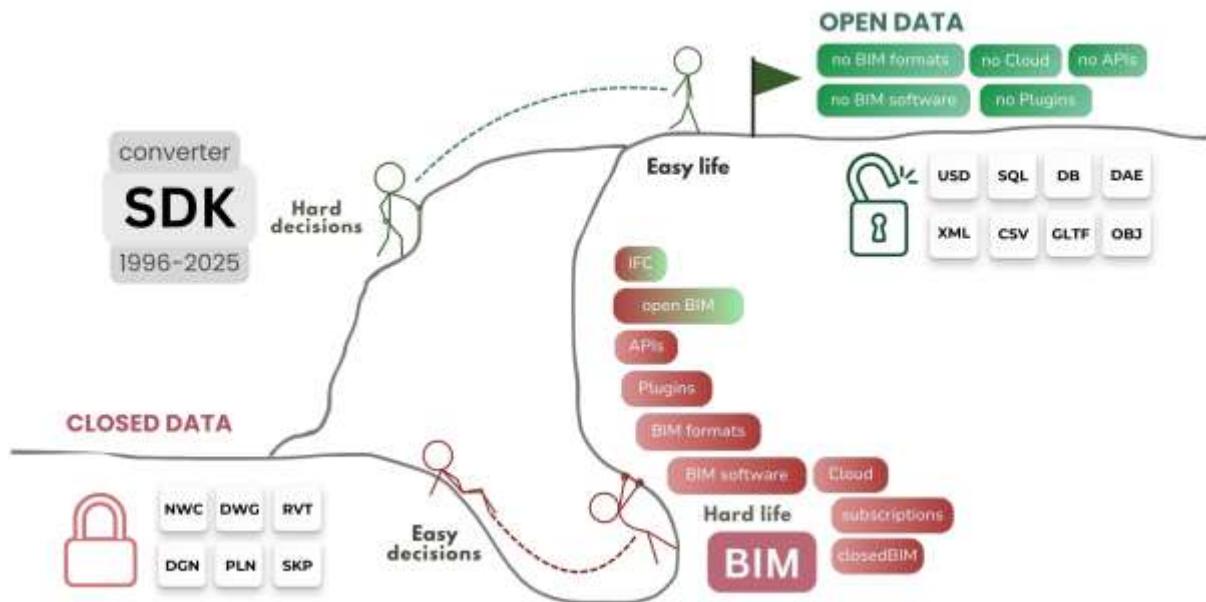


Attēls 6.4-1 Informācijas piesātinājuma cikls datu bāzēm biznesa procesiem būvniecības projektu īstenošanā.

Ja CAD (BIM) uzskatām par parametru pārsūtīšanas mehānismu atslēgu-vērtību pāru kopuma veidā, kas ģenerēts, pamatojoties uz ārpus projektēšanas vides definētām prasībām (6.4-1. attēls, 1.-2. posms), diskusijas uzmanības centrā ir nevis konkrēti programmatūras risinājumi un to ierobežojumi, bet gan fundamentālāki aspekti - datu struktūra, datu modeļi un datu prasības. Būtībā runa ir par datu bāzes parametru piesātinājumu un klasisko datu modelēšanas procesu (2.-3. un 5.-6. solis). Vienīgā atšķirība ir tāda, ka CAD datubāzu slēgtā rakstura un izmantoto formātu īpatnību dēļ šo procesu papildina specializētu BIM rīku izmantošana. Rodas jautājums: kāda ir BIM unikalitāte, ja citās nozarēs nav līdzīgu pieeju?

Pēdējo 20 gadu laikā BIM ir pozicionēts kā kas vairāk nekā tikai viens datu avots. CAD-BIM pakete bieži tiek reklamēta kā parametrisks rīks ar integrētu datu bāzi [64], kas spēj automatizēt projektēšanas, modelēšanas un būvniecības projektu dzīves cikla pārvaldības procesus. Tomēr realitātē BIM ir kļuvis vairāk par rīku, lai noturētu lietotājus pie piegādātāju platformas, nevis par ērtu datu un procesu pārvaldības metodi.

Rezultātā CAD- (BIM-) dati ir izolēti to platformās, slēpjot projekta informāciju aiz patentētiem API un ģeometrijas kodoliem. Tas ir atnēmis lietotājiem iespēju patstāvīgi piekļūt datubāzēm un iegūt, analizēt, automatizēt un pārsūtīt datus uz citām sistēmām, apejot piegādātāju ekosistēmas.



Attēls 6.4-2 Būvniecībā mūsdienā nepieciešami sarežģīti ģeometrijas kodoli, katru gadu atjaunināts API un īpašas licences CAD -(BIM-) programmatūrai.

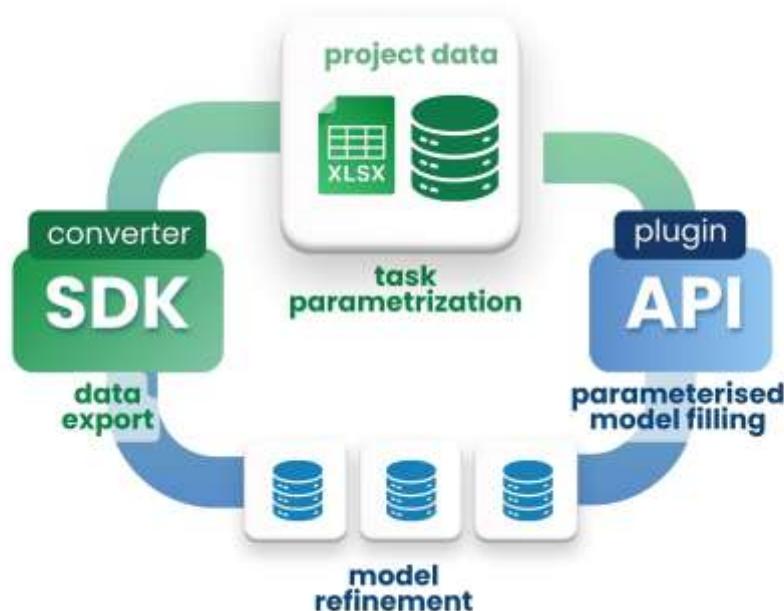
Uzņēmumiem, kas strādā ar mūsdienīgiem CAD rīkiem, būtu jāizmanto tāda pati pieeja darbam ar datiem, kādu praksē izmanto visi bez izņēmuma paši CAD pārdevēji: datu pārveidošana, izmantojot SDK - reversās inženierijas rīkus, pret ko CAD pārdevēji cīnās kopš 1995. gada [75]. Pilnībā piekļūstot CAD datu bāzei un izmantojot reversās inženierijas rīkus, mēs varam iegūt [127] plakānu vienību kopu ar atribūtiem un eksportēt tās jebkurā ērtā atvērtā formātā (6.4-2. attēls), ietverot gan ģeometriju, gan konstrukcijas elementu parametrus. Šī pieeja būtiski maina darba ar informāciju paradigmu - no uz failiem orientētas uz datiem orientētas arhitektūras:

- Tādi datu formāti kā RVT, IFC, PLN, DB1, CP2, CPIXML, USD, SQLite, XLSX, PARQUET un citi satur identisku informāciju par viena projekta elementiem. Tas nozīmē, ka zināšanām par konkrētu formātu un tā shēmu nevajadzētu būt šķērslim darbam ar pašiem datiem.
- Datus no jebkuriem formātiem var apvienot vienotā atvērtā strukturētā un granulārā struktūrā (9.1.1-10. attēls), kas satur MESH trīsstūrveida ģeometriju un visu objektu vienību īpašības bez ģeometrisko kodolu ierobežojumiem.
- Datu analīze tiecas pēc universāluma: izmantojot atvērtos datus, jūs varat strādāt ar projekta datiem neatkarīgi no izmantotā formāta.
- Minimizēta atkarība no API un piegādātāju spraudņiem: darbs ar datiem vairs nav atkarīgs no API prasmēm.

Kad un CAD -datu prasības tiek pārveidotas viegli analizējamos strukturētos attēlojuma formātos - izstrādātāji vairs nav atkarīgi no specifiskām datu shēmām un slēgtām ekosistēmām.

## Projektēšana, izmantojot parametrus: CAD nākotne un BIM

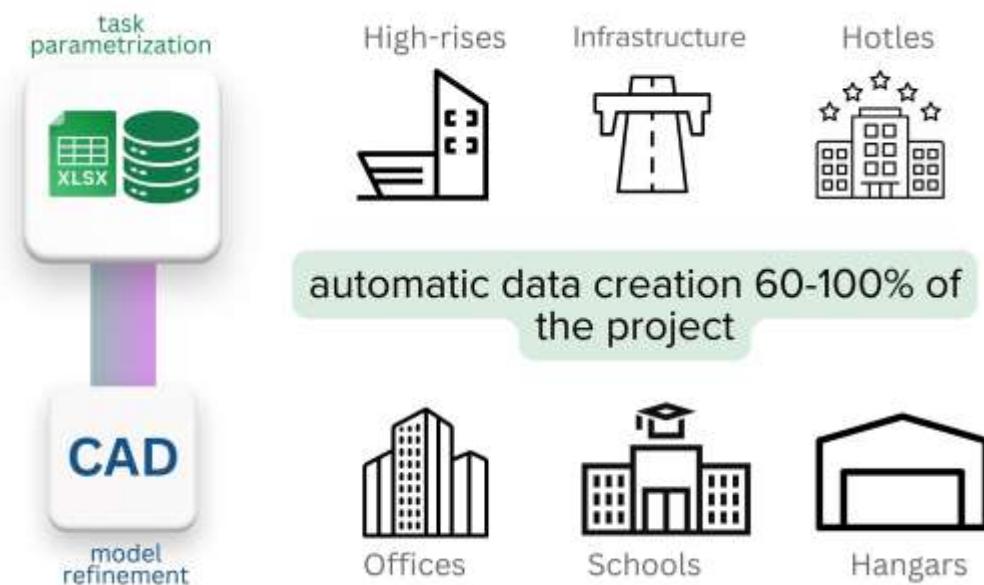
Neviens būvprojekts pasaulē vēl nekad nav sākts ar CAD programmu. Pirms rasējums vai modelis iegūst formu CAD programmā, tas iziet konceptualizācijas posmu ( 6.4-1. attēls, 1.-2. posms), kurā galvenā uzmanība tiek pievērsta parametriem, kas nosaka topošā objekta pamatideju un loģiku. Šis posms atbilst konceptuālajam līmenim datu modelēšanā ( 4.3-6. attēls). Parametri var pastāvēt tikai projektētāja prātā, bet ideālā gadījumā tie ir sakārtoti strukturētu sarakstu, tabulu veidā vai saglabāti datubāzēs ( 6.4-3. attēls), kas nodrošina pārrēdzamību, atkārtojamību un turpmāku projektēšanas procesa automatizāciju.



Attēls 6.4-3 Projektēšanas process ir iteratīvs process, kurā CAD datu bāze tiek papildināta ar informāciju no ārpuses, izmantojot vērtības kēdes prasības.

Pirms sākt pašu CAD modelēšanu (datu modelēšanas loģiskā un fiziskā fāze ( 4.3-7. attēls)), ir svarīgi noteikt robežparametru, kas kalpo par projekta pamatu. Šie raksturlielumi, tāpat kā citas prasības, tiek apkopoti pašā datu izmantošanas kēdes galā (piemēram, sistēmās), un ar to palīdzību jau tiek definēti ierobežojumi, mērķi un nākotnes projekta objektu galvenās īpašības.

Pašu modelēšanu var pilnībā automatizēt par 60-100%, izmantojot parametriskās modelēšanas rīkus ( 6.4-3. attēls), ja prasības ir labi definētas. Tiklīdz projekts ir aprakstīts parametru veidā, tā veidošana kļūst tehniski iespējama, piemēram, ar vizuālās programmēšanas valodu, piemēram, Grasshopper Dynamo, palīdzību, kas iestrādātas mūsdienu CAD -vidēs vai bezmaksas risinājumos Blender, UE, Omniverse.



Attēls 6.4-4 Lielākā daļa tipizēto projektu jau šodien tiek veidoti pilnīgi automātiski, pateicoties parametriskās programmēšanas rīkiem.

Jau mūsdienās lielus rūpnieciskus un tipveida projektus veido nevis dizaina nodaļa, bet gan parametriskie rīki un vizuālā programmēšana. Tas ļauj veidot modeli, pamatojoties uz datiem, nevis uz konkrēta dizainera vai vadītāja subjektīviem lēmumiem.

*Saturs ir svarīgāks par dizainu. Dizains bez satura nav dizains, bet gan dekorācija [128].*

- Džefrijs Zeldmans, tīmekļa dizainers un uzņēmējs

Process nesākas ar rasējumu vai 3D -modelēšanu, bet gan ar prasību formulēšanu. Tieši prasības nosaka, kādi elementi tiks izmantoti projektā, kādi dati jānodod citiem departamentiem un sistēmām. Tikai strukturētu prasību esamība ļauj automātiski regulāri (piemēram, pat ik pēc 10 minūtēm) pārbaudīt modeļus, nenovēršot projektētāja uzmanību no darba.

Iespējams, nākotnē CAD- (BIM-) sistēma kļūs tikai par saskarni datubāzes aizpildīšanai, un nebūs nozīmes, ar kādu CAD rīku veikta modelēšana (fiziskais līmenis).

Līdzīgi arī mašīnbūvē bieži tiek izmantota 3D modelēšana, taču tā nav obligāts vai obligāts projekta elements. Vairumā gadījumu pietiek ar klasisko 2D dokumentāciju, un uz tās pamata tiek izveidots nepieciešamais informācijas modelis. Šis modelis tiek veidots no komponentiem, kas strukturēti atbilstoši nozares standartiem, un satur visu nepieciešamo informāciju, lai izprastu projektēšanas un ražošanas organizāciju. Pēc tam no rūpīcas informācijas modeļa tiek izveidots rūpīcas informācijas modelis, kuram tiek pievienoti specifiski produkti un plūsmas diagrammas, kas jau ir orientētas uz tehnologu vajadzībām. Visu procesu var organizēt bez liekas sarežģītības, nepārslogojot sistēmu ar

3D grafikām, ja tās nesniedz reālas priekšrocības.

Ir svarīgi saprast, ka 3D modelim un CAD sistēmai nav jābūt galvenajai lomai - tas ir tikai kvantitatīvās un ģeometriskās analīzes instruments. Visi pārējie parametri, izņemot ģeometriju, kas raksturo objektu, ja iespējams, būtu jāuzglabā un jāapstrādā ārpus CAD vides (BIM).

Projektēšana, izmantojot parametrus, nav tikai tendenze, bet gan neizbēgama būvniecības nozares nākotne. Tā vietā, lai manuāli veidotu sarežģītus 3D -modeļus, projektētāji strādās ar datiem, apstiprinās tos un automatizēs procesus, tuvinot būvniecību programmēšanas pasaulei. Laika gaitā projektēšanas procesi tiks veidoti, pamatojoties uz programmatūras izstrādes principiem:

- Izveidot prasības → Izveidot modeli → Augšupielādēt uz serveri → Pārbaudīt izmaiņas → Pull request
- Pull pieprasījums automātiski veic modeļa pārbaudes attiecībā pret prasībām, kas tika izveidotas pirms projektēšanas procesa vai tā laikā.
- Pēc datu kvalitātes pārbaudes un apstiprināšanas izmaiņas tiek ieviestas projektā, kopējā datu bāzē vai automātiski pārnestas uz citām sistēmām.

Jau tagad mašīnbūvē šādas konstrukcijas izmaiņas sākas ar izmaiņu paziņojuma sagatavošanu. Līdzīga shēma gaidāma arī būvniecības nozarē: projektēšana būs iteratīvs process, kurā katrs solis tiks pamatots ar parametriskām prasībām. Šāda sistēma ļaus projektētājiem izveidot automatizētas pārbaudes un automatizētu pieprasījumu par konkrētu prasību izpildi.

Nākotnes dizaineris pirmkārt un galvenokārt ir datu operators, nevis manuāls modelētājs. Viņa uzdevums ir piepildīt projektu ar parametriskām vienībām, kur ģeometrija ir tikai viens no atribūtiem.

Tieši izpratnei par datu modelēšanas, klasifikācijas un standartizācijas nozīmi, kas tika detalizēti aplūkotas iepriekšējās grāmatas nodaļās, būs liela nozīme pārveidē. Nākotnes projektēšanas noteikumi tiks formalizēti kā atslēgas-vērtības parametru pāri XLSX vai XML -shēmu veidā.

Būvniecības nozares nākotne ir saistīta ar datu vākšanu, to analīzi, apstiprināšanu un procesu automatizāciju, izmantojot analītiskos rīkus. BIM (vai CAD) nav galīgais mērķis, bet tikai attīstības posms. Kad profesionāli sapratīs, ka viņi var strādāt tieši ar datiem, apejot tradicionālos CAD rīkus, pats termins "BIM" pakāpeniski atkāpsies no strukturētu un granulāru būvprojektu datu izmantošanas koncepcijām.

Viens no galvenajiem faktoriem, kas paātrina šo pārveidi, ir lielo valodas modeļu (LLM) un uz tiem balstītu rīku parādīšanās. Šīs tehnoloģijas maina veidu, kā tiek apstrādāti projektēšanas dati, ļaujot pieklūt informācijai bez nepieciešamības padzīlināti pārzināt API vai piegādātāju risinājumus. Izmantojot LLM, prasību izveides process un mijiedarbība ar CAD datiem kļūst intuitīva un pieejama.

## LLM parādīšanās projektēšanas CAD datu apstrādes procesos

Papildus CAD datubāzu piekļuves rīku un atvērtu un vienkāršotu CAD formātu izstrādei, LLM rīku (Large Language Models) parādīšanās revolucionizē projektēšanas datu apstrādi. Ja agrāk piekļuve informācijai galvenokārt tika nodrošināta, izmantojot sarežģītas saskarnes, un bija nepieciešamas programmēšanas prasmes un zināšanas par API, tad tagad ir iespējams mijiedarboties ar datiem, izmantojot dabisko valodu.

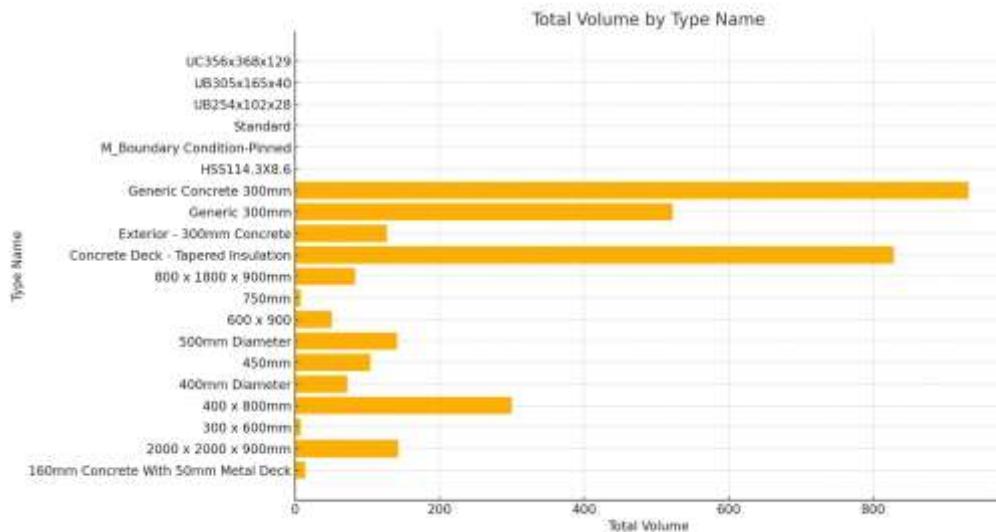
Inženieri, vadītāji un plānotāji bez tehniskas sagatavotības var iegūt nepieciešamo informāciju no projekta datiem, formulējot pieprasījumus parastā valodā. Ja dati ir strukturēti un pieejami (4.1-13.attēls), pietiek LLM tērzēšanā uzdot jautājumu, piemēram: "*Parādiet tabulā ar grupēšanu pēc tipa visas sienas, kuru tilpums ir lielāks par 10 kubikmetriem*", un modelis automātiski pārveidos šo pieprasījumu SQL vai kodā Pandas, ģenerējot kopsavilkuma tabulu, grafiku vai pat gatavu dokumentu.

Zemāk ir daži reāli piemēri, kā LLM -modeļi mijiedarbojas ar projektēšanas datiem, kas attēloti dažādos CAD- (BIM-) formātos.

- 🔗 Pieprasījuma piemērs LLM tērzēšanai uz CAD projektu RVT formātā pēc konvertēšanas (4.1-13.att.) tabulārajā datu rāmī (CHATGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkurā citā):

Grupējiet datus datu rāmī, kas iegūti no RVT faila, pēc "Type name" (Tipa nosaukums), summējot parametru "Volume" (Apjoms), un parādiet elementu skaitu grupā. Un, lūdzu, parādiet to visu kā horizontālu histogrammu bez nulles vērtībām.

- LLM atbildes reakcija kā horizontāla joslu diagramma (PNG formātā):



Attēls 6.4-5 . Tā vietā, lai, izmantojot spraudņus, veiktu 17 peles klikšķus vai ievadītu 40 koda rindas, LLM mēs uzreiz iegūstam QTO tabulu, izmantojot teksta vaicājumu.

- Lai izveidotu QTO tabulu ar sienu tipu kopplatību un daudzumu no kategorijas "Sienas", formulēsim teksta vaicājumu LLM -chat:

No projekta datu rāmja ņemiet tikai tos vienumus, kuru parametrā "Kategorija" ir "OST\_Walls", sagrupējiet tos pēc "Tipa nosaukums", saskaitiet slejas "Platība" vērtību, pieskaitiet daudzumu un parādiet tos tabulā, atņemot nulles vērtības.

- LLM atbilde pabeigtas QTO tabulas veidā:

Type Name	Total Area	Count
CL_W1	393.12 sq m	10
Cavity wall_sliders	9.37 sq m	1
Foundation - 300mm Concrete	30.90 sq m	1
Interior - 165 Partition (1-hr)	17.25 sq m	3
Interior - Partition	186.54 sq m	14
Retaining - 300mm Concrete	195.79 sq m	10
SH_Curtain wall	159.42 sq m	9
SIP 202mm Wall - conc clad	114.76 sq m	4
Wall - Timber Clad	162.91 sq m	8

**Attēls 6.4-6 QTO tabulas izveide dabīgajā valodā nodrošina tādas pašas kvalitātes rezultātu kā, izmantojot CAD - (BIM- ) rīkus.**

- Veidosim projekta vaicājumu IFC formātā pēc konvertēšanas tabulārajā datu rāmī un ievadīsim līdzīgu teksta vaicājumu jebkurā LLM tērzēšanā:

Nemiet no projekta tikai tos vienumus, kuriem parametrā "Vecāks" ir 1. un 2. līmenja vērtības, un nemiet vienumus, kuriem parametrā "Kategorija" ir IfcSlab vērtības, pēc tam sagrupējiet šos vienumus pēc parametra "Objekta tips", summējiet vērtības parametrā "PSet\_RVT \_Dimensions Area" un parādiet tās kā pīrāga diagrammu.

■ LLM atbilde kā pabeigta elementu grupu pīrāgu diagramma no IFC datiem :

For ■ Floor:127mm Slab on Grade, ■ Floor:150mm Exterior Slab on Grade,  
 ■ Floor:Finish Floor - Ceramic Tile, ■ Floor:Finish Floor - Wood, and  
 ■ Floor:Residential - Wood Joist with Subflooring



Attēls 6.4-7 IFC datu strukturēta formāta vaicājuma rezultāts var būt jebkura veida grafiks, kas ir ērts datu izpratnei.

Aiz katra no iegūtajiem gatavajiem risinājumiem ( 6.4-5. attēls - 6.4-7. attēls) slēpjās ducis Python koda rindu, izmantojot Pandas bibliotēku. Iegūto kodu var nokopēt no LLM tērzētavas un izmantot jebkurā vietējā vai tiesīsaistes IDE, lai iegūtu identiskus rezultātus ārpus LLM tērzētavas.

Tajā pašā LLM tērzētavā varam strādāt ne tikai ar projektiem, kas iegūti no 3D CAD (BIM) formātiem, bet arī ar plakaniem rasējumiem DWG formātā, kuriem varam veikt vaicājumus LLM tērzētavā, lai parādītu, piemēram, datus par elementu grupām līniju vai 3D ģeometrijas formā pēc konvertēšanas strukturētā formā.

## Automatizēta DWG failu analīze ar LLM un Pandas

DWG datu apstrādes process no DWG failiem informācijas nestrukturētības dēļ vienmēr ir bijis sarežģīts uzdevums, kas prasa specializētu programmatūru un bieži vien manuālu analīzi. Tomēr, attīstoties mākslīgajam intelektam un LLM rīkiem, ir kļuvis iespējams automatizēt daudzus posmus, kas mūsdienās lielākoties ir manuāls process. Aplūkosim reālu Pipeline pieprasījumu LLM (šajā piemērā ChatGPT), lai strādātu ar DWG rasējumiem, kas ļauj strādāt ar projektu:

- DWG datu filtrēšana pēc slāņa, ID un koordinātēm
- Vizualizējiet elementu ģeometriju
- Automātiski anotēt rasējumus, pamatojoties uz parametriem
- Sienu daudzliniju paplašināšana līdz horizontālajai plaknei

- Izveidojiet interaktīvu 3D - plakano datu vizualizāciju.
- Strukturēt un analizēt būvniecības datus bez sarežģītiem CAD -instrumentiem

Mūsu gadījumā cauruļvadu veidošanas process sākas ar secīgu koda ģenerēšanu, izmantojot LLM. Vispirms tiek ģenerēts uzdevumu aprakstošs vaicājums. ChatGPT ģenerē Python kodu, kas tiek izpildīts un analizēts, parādot rezultātu tērzēšanas telpā. Ja rezultāts neatbilst gaidītajam, pieprasījums tiek labots un process tiek atkārtots

**Cauruļvads** ir automatizētu darbību secība, kas tiek veikta, lai apstrādātu un analizētu datus. Šādā procesā katrs posms ievada datus, veic transformācijas un rezultātu nodod nākamajam posmam.

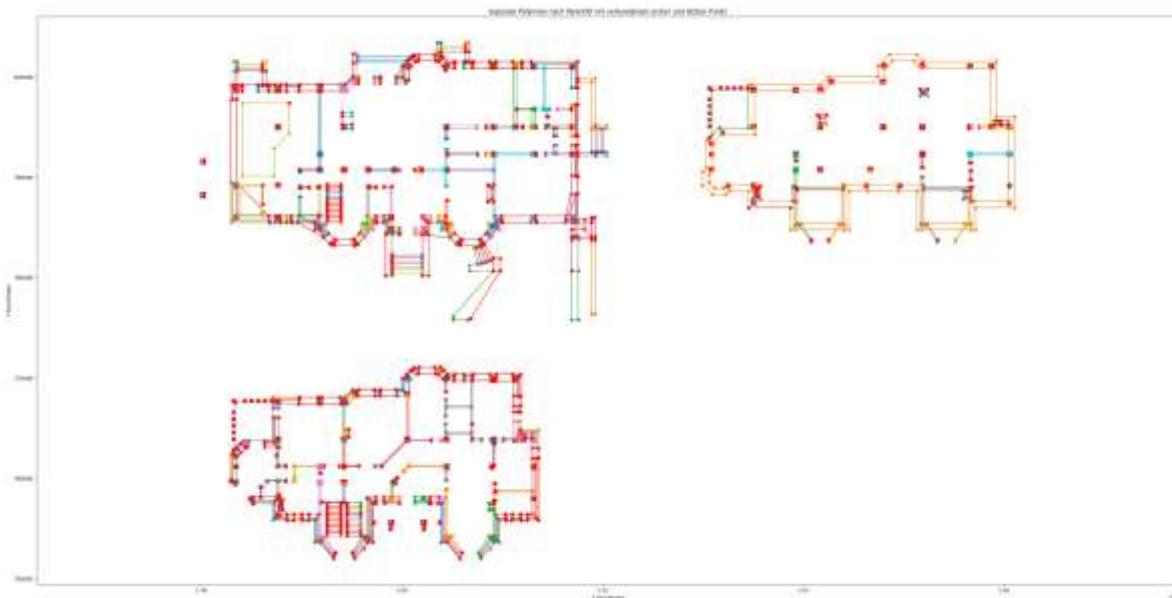
Pēc vēlamā rezultāta iegūšanas kods tiek nokopēts no LLM un bloku veidā ielīmēts kodā kādā no ērtajām IDE, mūsu gadījumā Kaggle platformā.com. Iegūtie koda fragmenti tiek apvienoti vienā cauruļvadā, kas automatizē visu procesu - no datu ielādes līdz to galīgajai analīzei. Šāda pieeja ļauj ātri izstrādāt un paplašināt analītiskos procesus bez dzīlām programmēšanas zināšanām. Visu turpmāk minēto fragmentu pilno kodu kopā ar parauga vaicājumiem var atrast Kaggle.com platformā, meklējot "DWG Analyse with ChatGPT | DataDrivenConstruction" [129]. [129].

Darba procesu ar DWG datiem pēc konvertēšanas strukturētā formā (4.1-13. attēls), sāksim ar klasisku soli - visu mūsu uzdevumam nepieciešamo rasējumu datu grupēšanu un filtrēšanu sienas elementiem, konkrēti polilīnijām (parametrs "ParentID" ļauj sagrupēt līnijas grupās), kuru parametrā (datu rāmja slejā) "Layer" ir virknes vērtība, kas satur šādu burtu kombināciju (RegEx) - "wall".

- Lai iegūtu kodu līdzīgam uzdevumam un rezultātu attēla veidā, jums jāraksta šāds pieprasījums LLM:

Vispirms pārbaudiet, vai no DWG iegūtajā datu rāmī ir norādītās kolonas: "Layer", "ID", "ParentID" un "Point". Pēc tam izfiltrējet ID no slejas "Layer", kas satur virknī "wall" (siena). Slejā "ParentID" atrodiet vienumus, kas atbilst šiem identifikatoriem. Definējiet funkciju, lai attīrītu un sadalītu datus slejā "Point". Tas ietver iekavju nonemšanu un vērtību sadalīšanu "x", "y" un "z" koordinātēs. Uzzīmējiet datus, izmantojot matplotlib. Katram unikālajam "ParentID" uzzīmējiet atsevišķu daudzliniju, kas savieno "Point" koordinātas. Ja iespējams, pārliecinieties, ka pirmsais un pēdējais punkts ir savienoti. Iestatiet atbilstošus marķējumus un nosaukumus, nodrošinot, ka x un y asis ir vienādā mērogā.

- Atbildi LLM sniegs jums gatavu attēlu, aiz kura slēpjās to ģenerējošais Python kods:



Attēls 6.4-8 LLM kods no DWG faila ieguva visas "sienas" slāņa līnijas, noskaidroja to koordinātas un konstruēja daudzlinijas, izmantojot vienu no Python bibliotēkām.

- Tagad pievienosim līnijām platības parametru, kas katrai polilīnijai ir norādīts tās īpašībās (vienā no datu rāmja kolonnām):

Tagad iegūstiet tikai vienu "ParentID" no katras polilīnijas - atrodiet šo ID kolonnā "ID", ņemiet "Area" vērtību, daliet ar 1 000 000 un pievienojiet šo vērtību diagrammai.

- LLM atbilde parādīs jaunu grafiku, kurā katrai daudzlinijai būs uzraksts ar tās platību:

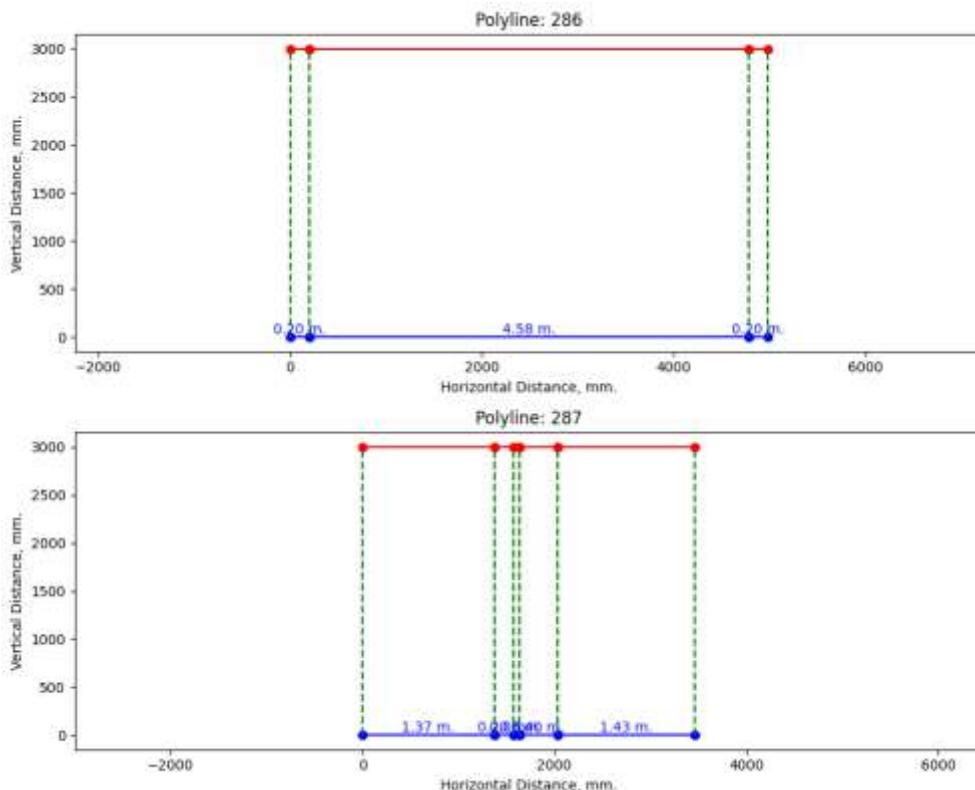


**Attēlā 6.4-9** LLM ir pievienots kods, kas nem platības vērtības katrai daudzlinijai un pievieno tās attēlam ar līniju vizualizāciju.

- Pēc tam katru daudzliniju pārveidosim par horizontālu līniju, pievienosim paralēlu līniju 3000 mm augstumā un savienosim tās vienā plaknē, lai šādā veidā parādītu sienas elementu virsmu izkārtojumu:

Jums ir jāņem visi elementi no slejas "Layer" ar vērtību "wall". Paņemiet šos ID kā sarakstu no slejas "ID" un atrodiet šos ID no visa datu rāmja slejā "ParentID". Visi elementi ir līnijas, kas ir apvienotas vienā polilīnijā. Katrai līnijai ir atšķirīga pirmā punkta x, y ģeometrija slejā "Point" (punkts). Katra polilīnija ir jāņem pēc kārtas un no punkta 0,0 horizontāli jānozīmē katra polilīnijas segmenta garums. katras polilīnijas segmenta garums vienā līnijā. Pēc tam uzzīmējiet tieši tādas pašas līnijas, tikai 3000 augstākas, savienojiet visus punktus vienā plaknē.

- ❓ LLM atbilde izvadīs kodu, kas ļauj uzzīmēt sienu rasējumus plaknē:

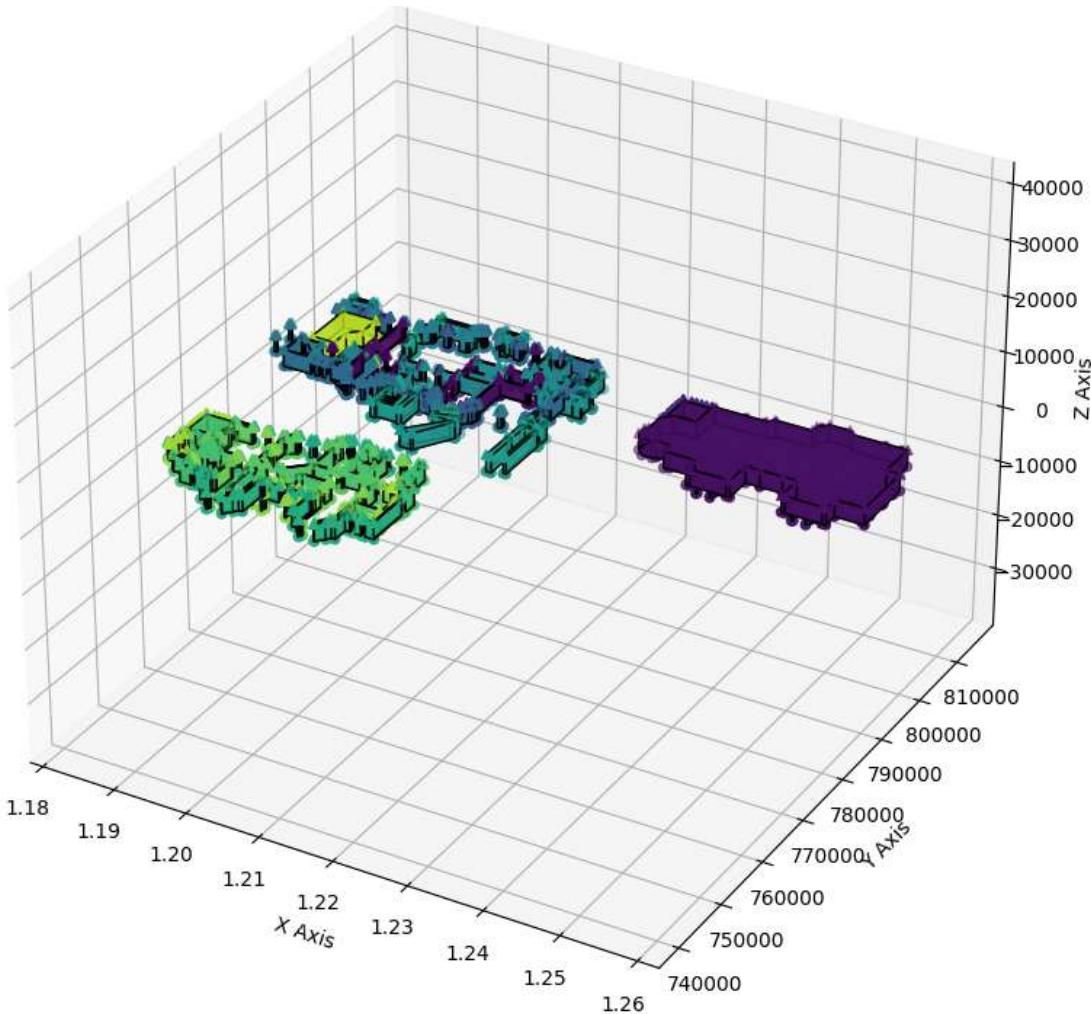


**Attēlā 6.4-10.** Katru daudzliniju, izmantojot uzvednes, pārvēršam maketā, kas vizualizē sienu plaknes tieši LLM tērzēšanas logā.

- ❓ Tagad pāriesim no 2D projekcijas uz 3D - modelēsim sienas no plakanām līnijām, savienojot augšējo un apakšējo daudzliniju slāni:

Vizualizējet sienas elementus 3D vietnē, savienojot daudzlinijas  $z = 0$  un  $z = 3000$  mm augstumā. Lai izveidotu slēgtu ģeometriju, kas atveido ēkas sienas. Izmantojet Matplotlib 3D grafēšanas rīku.

- LLM izveidos interaktīvu 3D grafiku, kurā katra polilīnija tiks attēlota kā plakņu kopums. Lietotājs varēs brīvi pārvietoties starp elementiem ar datora peli, izpētot modeli 3D režīmā, kopējot kodu no tērzēšanas uz IDE:



Attēlā 6.4-11 LLM palīdzēja izveidot kodu [129], lai vizualizētu plakanas rasēšanas līnijas 3D attēlā, ko var izpētīt IDE vidē esošajā 3D skatītājā.

Lai izveidotu loģisku un reproducējamu cauruļvadu - no sākotnējās DWG -faila konvertēšanas un ielādēšanas līdz galarezultātam -, ieteicams pēc katras soļa kopēt ģenerēto LLM -koda bloku uz IDE. Šādā veidā jūs ne tikai pārbaudīsiet rezultātu tērzēšanas laikā, bet arī uzreiz palaidīsiet to savā izstrādes vidē. This allows you to build the process sequentially, debugging and adapting it as needed.

Pilnu visu fragmentu cauruļvada kodu (attēli 6.4-8 līdz attēls 6.4-11) kopā ar paraugu pieprasījumiem var atrast Kaggle platformā.com, meklējot "DWG Analyse with ChatGPT | DataDrivenConstruction". [129]. Kaggle vietnē var ne tikai apskatīt kodu un izmantotos uzvedījus, bet arī bez maksas kopēt un

testēt visu cauruļvadu ar oriģinālajiem DWG datu rāmjiem mākonī, neinstalējot papildu programmatūru vai pašu IDE.

Šajā nodaļā aprakstītā pieeja ļauj pilnībā automatizēt dokumentu pārbaudi, apstrādi un ģenerēšanu, pamatojoties uz DWG -projektiem. Izstrādātā cauruļvadu sistēma ir piemērota gan atsevišķu rasējumu apstrādei, gan desmitiem, simtiem un tūkstošiem DWG failu apstrādei partijās, automātiski ģenerējot nepieciešamos pārskatus un vizualizācijas katram projektam.

Procesu var organizēt secīgi un pārskatāmi: vispirms dati no CAD datnes tiek automātiski konvertēti XLSX formātā, pēc tam ielādēti datu rāmī, kam seko grupēšana, pārbaude un rezultātu ģenerēšana - tas viss tiek īstenots ar vienu Jupyter piezīmjdatoru vai Python -skriptu jebkurā populārā IDE. Vajadzības gadījumā procesu var viegli paplašināt, integrējot ar projekta dokumentācijas pārvaldības sistēmām: CAD failus var automātiski iegūt atbilstoši noteiktiem kritērijiem, rezultātus var atgriezt atpakaļ glabāšanas sistēmā, un lietotājiem var paziņot, kad rezultāti ir gatavi, - pa e-pastu vai ziņojumapmaiņu.

Izmantojot LLM čatus un aģentus darbam ar projektēšanas datiem, tiek samazināta atkarība no specializētām CAD programmām un iespējams veikt arhitektūras projektu analīzi un vizualizāciju bez nepieciešamības manuāli mijiedarboties ar interfeisu - bez peles klikšķiem un sarežģītas izvēlēju navigācijas atcerēšanās.

Ar katru dienu būvniecības nozarē arvien vairāk dzirdēsim par LLM, granulāriem strukturētiem datiem, DataFrames un kolonnveida datu bāzēm. Vienotas divdimensiju DataFrames, kas veidotas no dažādām datu bāzēm un CAD formātiem, būs ideāla degviela moderniem analītiskiem rīkiem, ar kuriem aktīvi strādā citu nozaru speciālisti.

Pats automatizācijas process tiks ievērojami vienkāršots - tā vietā, lai studētu slēgtu nišas produktu API un rakstītu sarežģītus skriptus parametru analīzei vai pārveidošanai, tagad pietiks formulēt uzdevumu atsevišķu teksta komandu kopuma veidā, kas tiks salocīts vajadzīgajā cauruļvadā vai darba plūsmas procesā vajadzīgajā programmēšanas valodā, kas darbojas bez maksas gandrīz jebkurā ierīcē. Nebūs jāgaida jauni produkti, formāti, spraudņi vai atjauninājumi no CAD- (BIM-) rīku piegādātājiem. Inženieri un būvnieki varēs patstāvīgi strādāt ar datiem, izmantojot vienkāršus, bezmaksas un viegli saprotamus rīkus, kuriem palīdzēs LLM tērzēšanas un aģenti.

## Nākamie soļi: pāreja no slēgtiem formātiem uz atvērtiem datiem

Strādājot ar nākotnes projektēšanas datiem, diez vai kādam patiešām būs jāizprot patentētu rīku ģeometriskie kodoli vai jāapgūst simtiem nesaderīgu formātu, kas satur vienu un to pašu informāciju. Tomēr, nesaprodot, kāpēc pāreja uz atvērtiem strukturētiem datiem ir svarīga, ir grūti argumentēt jaunu bezmaksas rīku, atvērtu datu un pieeju izmantošanu, ko, visticamāk, neveicinās programmatūras ražotāji.

Šajā nodaļā mēs esam aplūkojuši CAD (BIM) datu galvenās iezīmes, to ierobežojumus un iespējas, kā arī to, ka, neraugoties uz ražotāju mārketinga solījumiem, inženieri un projektētāji ik dienu saskaras ar grūtībām, iegūstot, pārsūtot un analizējot projektēšanas informāciju. Izpratne par šo sistēmu

arhitektūru un iepazīšanās ar alternatīvām pieejām, kas balstītas uz atvērtiem formātiem un automatizāciju, izmantojot LLM, var ievērojami atvieglot dzīvi pat vienam profesionālim, nemaz nerunājot par uzņēmumiem. Apkopojoši šo daļu, ir vērts izcelt galvenos praktiskos soļus, kas palīdzēs jums pielietot šajā sadaļā aplūkotās pieejas jūsu ikdienas uzdevumos:

- Paplašiniet savu rīku klāstu darbam ar projekta datiem
  - Izpētiet pieejamos spraudņus un paīgprogrammas, lai iegūtu datus no jūsu izmantotajām CAD - (BIM-) sistēmām.
  - Izpētiet pieejamos SDK un API, kas ļauj automatizēt datu ieguvi no slēgtiem formātiem bez nepieciešamības manuāli atvērt specializētu programmatūru.
  - Apgūt pamatprasmes darbā ar atvērtajiem neparametriskās ģeometrijas formātiem (OBJ, glTF, USD, DAE) un attiecīgajām atvērtā koda bibliotēkām.
  - Mēģiniet izdomāt sistēmu projekta metadatu glabāšanai atsevišķi no ģeometrijas ārpus CAD (BIM) risinājumiem, lai vienkāršotu analīzi un integrāciju ar citām sistēmām.
  - Izmantojiet LLM, lai automatizētu datu konvertēšanu starp formātiem.
- Izveidojiet savus procesus projekta informācijas apstrādei
  - sākt aprakstīt uzdevumus un modelēšanas prasības, izmantojot parametrus un to vērtības vienkāršos un strukturētos formātos.
  - Izveidojiet personīgo skriptu vai koda bloku bibliotēku bieži veicamām operācijām.
- Veiciniet atvērto standartu izmantošanu savā darbā.
  - Aicināt kolēģus un partnerus dalīties ar datiem atvērtos formātos, kurus neierobežo programmatūras piegādātāju ekosistēma.
  - Ar konkrētiem piemēriem demonstrēt strukturētu datu izmantošanas priekšrocības.
  - Iniciēt diskusijas par slēgto formātu problēmām un iespējamiem risinājumiem.

Pat ja jūs nevarat mainīt sava uzņēmuma politiku attiecībā uz CAD - (BIM-) platformām, personīga izpratne par principiem, kā strādāt ar projekta datiem atvērtos formātos, ļaus jums ievērojami palielināt sava darba efektivitāti. Izveidojot savus rīkus un metodes datu ieguvei un pārveidošanai no dažādiem formātiem, jūs ne tikai optimizēsiet savu darba plūsmu, bet arī iegūsiet elastību, lai apietu standarta programmatūras risinājumu ierobežojumus.



## VII DALĀ

### **UZ DATIEM BALSTĪTA LĒMUMU PIENEMŠANA, ANALĪTIKA, AUTOMATIZĀCIJA UN MAŠĪNMĀCĪŠANĀS.**

Septītajā daļā uzmanība pievērsta datu analīzei un procesu automatizācijai būvniecības nozarē. Tajā ir aplūkots, kā dati kļūst par pamatu lēmumu pieņemšanai, un izskaidroti informācijas vizualizācijas principi efektīvai analīzei. Detalizēti aprakstīti galvenie rezultatīvie rādītāji (KPI), ieguldījumu atdeves novērtēšanas metodes (ROI) un informācijas paneļu izveide projektu uzraudzībai. Īpaša uzmanība ir pievērsta ETL procesiem (Extract, Transform, Load) un to automatizācijai, izmantojot cauruļvadus (Pipeline), lai pārveidotu atšķirīgus datus strukturētā informācijā analīzei. Apskatīti tādi darbplūsmas orkestrēšanas rīki kā Apache Airflow, Apache NiFi un n8n, kas ļauj veidot automatizētus datu cauruļvadus bez dziļām programmēšanas zināšanām. Nozīmīga loma ir lielajiem valodas modeļiem (LLM) un to izmantošanai, lai vienkāršotu datu analīzi un automatizētu rutīnas uzdevumus.

## NODĀLA 7.1.

### DATU ANALĪTIKA UN UZ DATIEM BALSTĪTU LĒMUMU PIENEMŠANA.

Pēc informācijas apkopošanas, strukturēšanas, attīrišanas un pārbaudes posmiem ir izveidots saskaņots un analizējams datu kopums. Iepriekšējās grāmatas daļas tika sistematizēti un strukturēti neviendabīgi avoti - no PDF dokumentiem un sanāksmju teksta ierakstiem līdz CAD modeļiem un ģeometriskiem datiem. Detalizēti aprakstīts informācijas pārbaudes un saskaņošanas process ar dažādu sistēmu un klasifikatoru prasībām, novēršot dublēšanos un neatbilstības.

Visi ar šiem datiem veiktie aprēķini (trešā, ceturtā grāmatas daļa) - no vienkāršām transformācijām līdz laika, izmaksu un ESG rādītāju aprēķiniem (piektā daļa) - ir apkopoti analīzes uzdevumi. Tie veido pamatu, lai izprastu projekta pašreizējo stāvokli, novērtētu tā parametrus un pēc tam pieņemtu lēmumus. Rezultātā dati aprēķinu rezultātā no atšķirīgu ierakstu kopuma pārvēršas par pārvaldāmu resursu, kas spēj atbildēt uz galvenajiem uzņēmējdarbības jautājumiem.

Iepriekšējās nodaļās ir sīki izklāstīti datu vākšanas un kvalitātes kontroles procesi, kas paredzēti izmantošanai tipiskos uzņēmējdarbības gadījumos un būvniecības nozarei raksturīgos procesos. Analīтика šajā kontekstā daudzējādā ziņā ir līdzīga lietojumiem citās nozarēs, taču tai ir vairākas specifiskas iezīmes.

Turpmākajās nodaļās tiks detalizēti aprakstīts datu analīzes process, tostarp automatizācijas posmi no sākotnējās informācijas iegūšanas un tās pārveidošanas līdz tās turpmākai pārsūtīšanai uz mērķa sistēmām un dokumentiem. Vispirms tiks izklāstīta teorētiskā daļa, pievēršoties atsevišķiem uz datiem balstītas lēmumu pieņemšanas aspektiem. Pēc tam turpmākajās nodaļās tiks izklāstīta praktiskā daļa, kas saistīta ar ETL automatizāciju un konstruēšanu -Pipeline.

#### **Dati kā resursu lēmumu pieņemšanā**

Uz datiem balstīta lēmumu pieņemšana bieži ir iteratīvs process, kas sākas ar sistemātisku informācijas vākšanu no dažādiem informācijas avotiem. Līdzīgi kā dabas ciklā, atsevišķi datu elementi un veselas informācijas sistēmas pakāpeniski nonāk augsnē - uzkrājas uzņēmumu informācijas krātuvēs (1.3-2. attēls). Laika gaitā šie dati, līdzīgi kā kritušās lapas un zari, pārvēršas vērtīgā materiālā. Datu inženieru un analītiķu micēlijs organizē un sagatavo informāciju turpmākai izmantošanai un pārvērš kritušos datus un sistēmas vērtīgā kompostā, lai izaugtu jauni dzinumi un jaunas sistēmas (1.2-5. attēls).

Tendencies plaši izmantot analītiku dažādās nozarēs iezīmē jauna laikmeta sākumu, kad darbs ar datiem klūst par profesionālās darbības pamatu (7.1-1. attēls). Būvniecības nozares profesionāļiem ir svarīgi pielāgoties šīm pārmaiņām un būt gataviem pāriet jaunā laikmetā - datu un analītikas laikmetā

Manuāla datu pārvietošana starp tabulām un manuāla aprēķinu veikšana pamazām paliek pagātnē, dodot vietu automatizācijai, datu plūsmas analīzei, analīzei un mašīnmācībai. Šie rīki klūst par mūsdienu lēmumu atbalsta sistēmu pamatelementiem.

McKinsey grāmatā "Rebooting. McKinsey's Guide to Overcoming Competition in the Age of Digital Technology and Artificial Intelligence" [130] citēts 2022. gadā veikts 1330 augstākā līmeņa vadītāju no dažādiem reģioniem, nozarēm un funkcionālajām jomām apsekojums. [130] citēts 2022. gadā veikts pētījums, kurā piedalījās 1330 augstākā līmeņa vadītāji no dažādiem reģioniem, nozarēm un funkcionālajām jomām. Saskaņā ar tā rezultātiem 70% vadītāju izmanto progresīvu analītiku, lai ġenerētu savas idejas, un 50% ievieš mākslīgo intelektu, lai uzlabotu un automatizētu lēmumu pieņemšanas procesus.



Attēls 7.1-1 Datu analīze un analītika ir galvenais rīks, lai palielinātu lēmumu pieņemšanas ātrumu uzņēmumā.

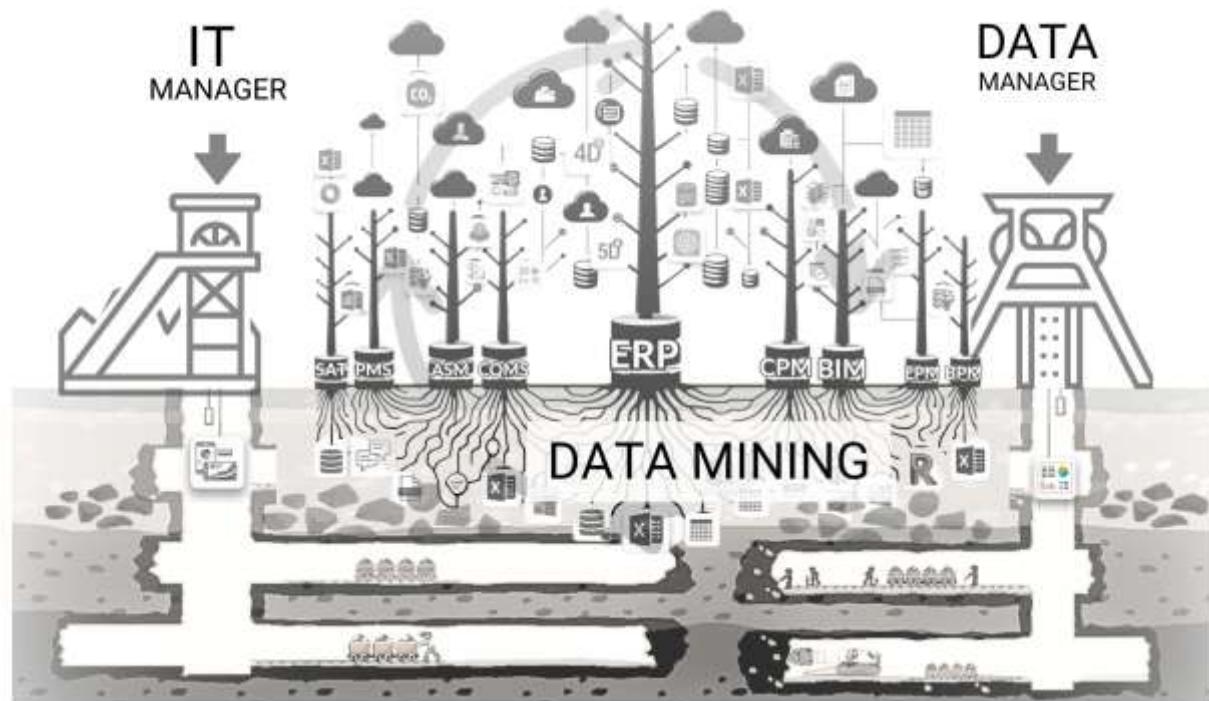
Datu analīze, līdzīgi kā izplatās micēlijs, iekļūst pagātnes lēmumu humusā, palīdzot savienot atsevišķas sistēmas un virzot vadītājus uz vērtīgām atziņām. Šīs zināšanas, līdzīgi kā barības vielas no sapuvušiem datu sistēmas kokiem, baro jaunus lēmumus uzņēmumā, novedot pie efektīvām pārmaiņām un kvalitatīvas informācijas augšanas, līdzīgi kā jauni dzinumi un asni, kas rodas no bagātas un veselīgas augsnes (1.2-5. attēls).

*Skaitļiem ir svarīgs stāsts. Tie paļaujas uz to, ka jūs tiem piešķirsiet skaidru un pārliecinošu balsi [131].*

- Stephen Few, datu vizualizācijas eksperts

Vidēja lieluma un mazos uzņēmumos informācijas ieguve un sagatavošana turpmākai analīzei mūsdienās ir ārkārtīgi darbielīgs process (7.1-2. attēls), ko var salīdzināt ar astoņpadsmitā gadsimta oglūpniecību. Vēl nesen datu ieguves un sagatavošanas darbs bija drīzāk rezervēts piedzīvojumu meklētājiem, kas strādā šauri specializētā nišā ar nelielu un ierobežotu rīku kopumu darbam ar dažāda veida datiem no nestukturētiem, brīvi strukturētiem, jauktiem un slēgtiem avotiem.

Lēmumu pieņēmējiem un vadītājiem bieži vien nav pieredzes darbā ar heterogēniem datiem un sistēmām, tomēr viņiem ir jāpieņem uz tiem balstīti lēmumi. Rezultātā uz datiem balstīta lēmumu pieņemšana mūsdienu būvniecības nozarē pēdējās desmitgadēs ir mazāk līdzinājusies automatizētam procesam un vairāk līdzinājusies daudzu dienu manuālajam kalnrača darbam agrīnajās oglraktuvēs.



**Attēls 7.1-2 Datu ieguves procesā speciālisti veic sarežģītu datu sagatavošanas ceļu - no attīrišanas līdz strukturēšanai turpmākai analīzei.**

Lai gan mūsdienu datu ieguves metodes būvniecības nozarē noteikti ir modernākas nekā 12. gadsimta kalnraču primitīvie paņēmieni, tas joprojām ir sarežģīts un riskants uzdevums, kas prasa ievērojamus resursus un zināšanas, ko var atļauties tikai lieli uzņēmumi. Datu ieguves un analīzes procesus no iepriekšējo projektu uzkrātā mantojuma līdz šim pārsvarā veica lieli, tehnoloģiski attīstīti uzņēmumi, kas datus ir konsekventi vākuši un glabājuši gadu desmitiem

Agrāk vadošā loma analītikā bija tehnoloģiski nobriedušiem uzņēmumiem, kas datus bija krājuši gadu desmitiem. Šodien situācija mainās: piekļuve datiem un datu apstrādes rīkiem kļūst demokrātiskāka - iepriekš sarežģīti risinājumi tagad ir pieejami ikvienam bez maksas.

Analītikas izmantošana ļauj uzņēmumiem pieņemt precīzākus un pamatošākus lēmumus reālajā laikā. Tālāk sniegtais gadījuma pētījums ilustrē, kā vēsturiskie dati var palīdzēt pieņemt finansiāli pamatotus lēmumus:

- ➲ **Projekta vadītājs** - "Tagad vidējā betona cena pilsētā ir 82€ /m<sup>3</sup>, mūsu tāmē ir 95 €/m<sup>3</sup>."
- ➲ **Aplēšu sastādītājs** - "Iepriekšējos projektos pārsniegums bija aptuveni 15%, tāpēc es atteicos."
- ➲ **Datu pārvaldnieks vai klientu puses kontroles inženieris** - "Apskatīsim pēdējo trīs konkursu analītiskos datus."

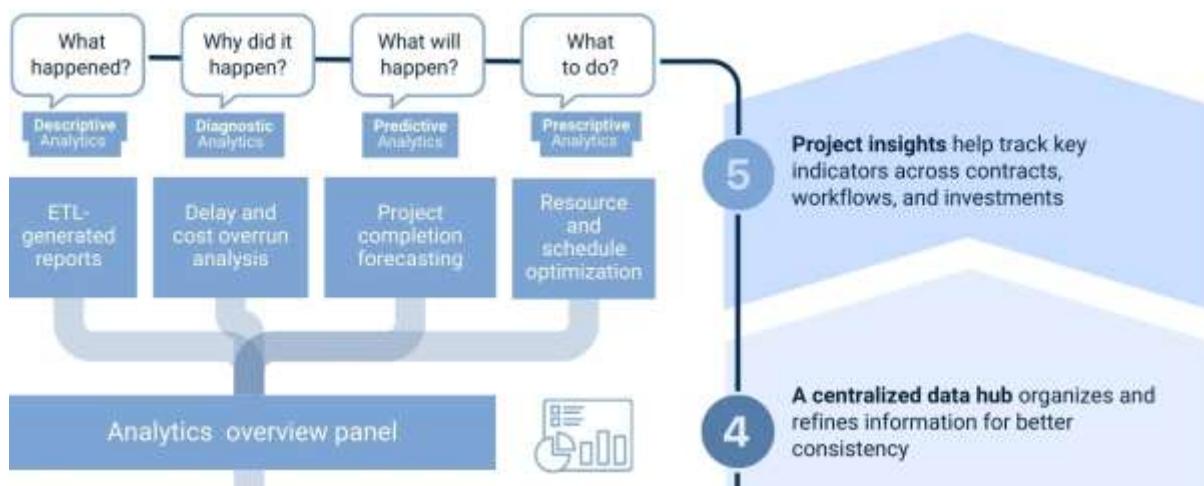
Analizējot iepriekšējo projektu DataFrame, mēs iegūstam:

- **Vidējā faktiskā iepirkuma cena: 84,80 /m<sup>3</sup>**

■ Vidējais pārtērēto līdzekļu īpatsvars: +4,7%.

■ Ieteicamā likme aprēķinā: ~ 85 /m<sup>3</sup>€

Šāds lēmums vairs nebūs balstīts uz subjektīvām izjūtām, bet gan uz konkrētu vēsturisku statistiku, kas palīdz samazināt riskus un palielināt konkursa piedāvājuma derīgumu. Iepriekšējo projektu datu analīze kļūst par sava veida "organisko mēslojumu", no kura dīgst jauni, precīzāki risinājumi.



Attēls 7.1-3 Datu analīze sniedz atbildes uz trim galvenajiem jautājumiem: kas notika, kāpēc tas notika un kas jādara tālāk.

Lēmumu pieņemēji un vadītāji bieži saskaras ar nepieciešamību strādāt ar neviendabīgiem datiem un sistēmām bez pietiekamām tehniskām zināšanām. Šādās situācijās vizualizācija, kas ir viens no pirmajiem un vissvarīgākajiem soļiem analītiskajā procesā, ir galvenais palīgs datu izpratnei. Tā ļauj informāciju pasniegt vizuālā un saprotamā veidā.

### datu vizualizācija: izpratnes un lēmumu pieņemšanas atslēga

Mūsdienu būvniecības nozarē, kur projektu datus raksturo sarežģītā un daudzlīmenī struktūra, vizualizācijai ir būtiska nozīme. Datu vizualizācija ļauj projektu vadītājiem un inženieriem vizualizēt sarežģītus modeļus un tendencies, kas slēpjas lielos, neviendabīgos datu apjomos.

Vizualizējot datus, ir vieglāk saprast projekta statusu: resursu sadalījumu, izmaksu tendencies vai materiālu izmantošanu. Grafiki un diagrammas padara sarežģītu un sausu informāciju pieejamu un saprotamu, ļaujot ātri noteikt galvenās jomas, kurām jāpievērš uzmanība, un pamanīt iespējamās problēmas.

Datu vizualizācija ne tikai atvieglo informācijas interpretāciju, bet arī ir būtisks solis analītiskajā procesā un pamatojuši vadības lēmumu pieņemšanā, palīdzot atbildēt uz jautājumiem "kas notika?" un "kā tas notika?" (2.2-5. attēls). (2.2-5. attēls).

Grafikas ir vizuāli rīki logisko problēmu risināšanai [132].

- Jacques Bertin, "Grafika un grafiskā informācijas apstrāde"

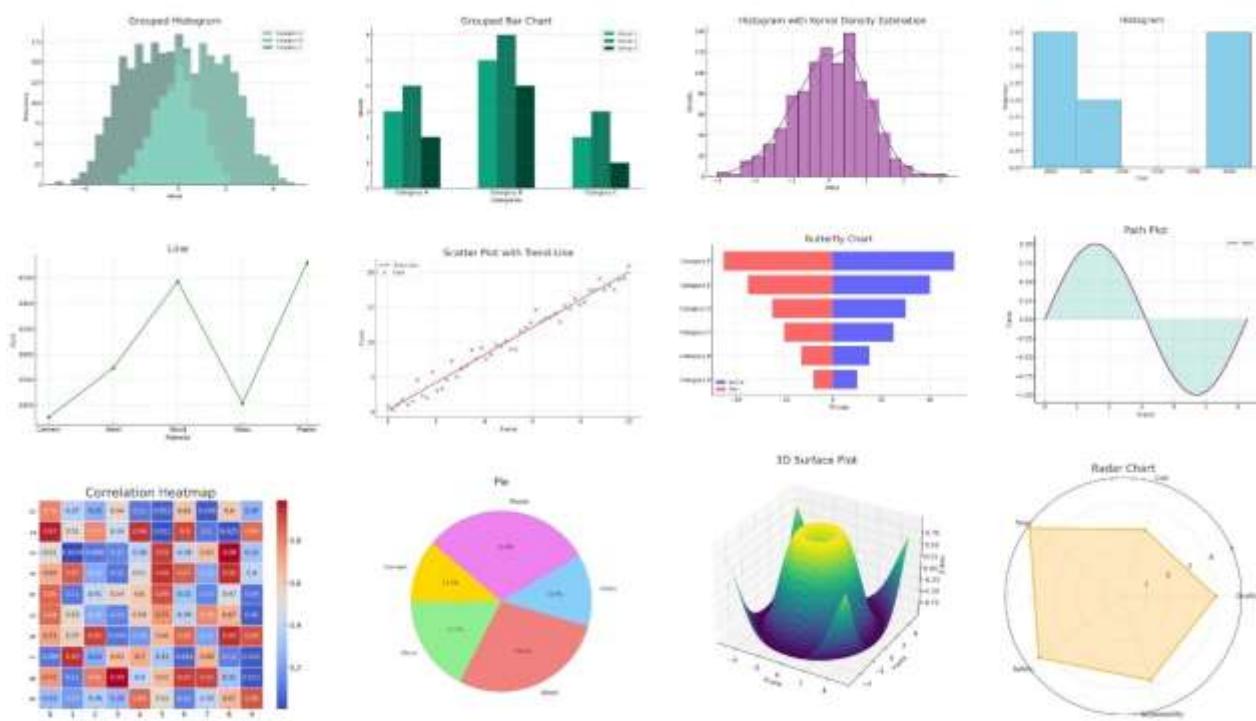
Pirms svarīgu lēmumu pieņemšanas projektu vadītāji biežāk izmanto vizuālu datu attēlojumu, nevis sausus un grūti interpretējamus skaitļus no izklājlapām vai teksta ziņojumiem.

Dati bez vizualizācijas ir kā būvlaukumā nejauši izkaisīti būvmateriāli - to potenciāls ir neskaidrs. Tikai tad, kad tie tiek skaidri vizualizēti, kā no ķieģeljiem un betona būvēta māja, kļūst skaidra to vērtība. Kamēr māja nav uzbūvēta, nav iespējams pateikt, vai materiālu kaudze kļūs par mazu mājiņu, greznu villu vai debesskrāpi.

Uzņēmumu rīcībā ir dati no dažādām sistēmām (no 1.2-4. attēla līdz 2.1-10. attēlam), finanšu darījumi un plaši teksta dati. Tomēr šo datu izmantošana uzņēmējdarbības labā bieži vien ir sarežģīta. Šādās situācijās vizualizācija kļūst par svarīgu rīku datu nozīmes nodošanai, palīdzot sniegt informāciju jebkuram ekspertam saprotamā formātā, piemēram, paneļos, grafikos un diagrammās.

PwC pētījumā "What Students Need to Succeed in a Rapidly Changing Business World" (2015) [9] uzsvērts, ka veiksmīgi uzņēmumi neaprobežojas tikai ar datu analīzi, bet aktīvi izmanto interaktīvus vizualizācijas rīkus, piemēram, grafikus, infografikas un analītiskus paneļus, lai atbalstītu lēmumu pieņemšanu. Saskaņā ar ziņojumu - datu vizualizācija palīdz klientiem izprast stāstu, ko stāsta dati, izmantojot diagrammas, grafikus, diagrammas, informācijas paneļus un interaktīvus datu modeļus.

Informācijas pārveidošana vizuālās grafiskās formās, piemēram, diagrammās, grafikos un diagrammās, uzlabo cilvēka smadzeņu izpratni un datu interpretāciju ( 7.1-4. attēls). Tas ļauj projektu vadītājiem un analītiķiem ātrāk novērtēt sarežģītus scenārijus un pieņemt pamatotus lēmumus, pamatojoties uz vizuāli atpazīstamām tendencēm un modeļiem, nevis intuīciju.



Attēls 7.1-4 Dažāda veida vizualizācija ir izstrādāta, lai palīdzētu cilvēka smadzenēm labāk saprast un izprast sauso skaitļu informāciju.

Vizualizāciju izveide no datiem un dažādu bezmaksas vizualizācijas bibliotēku izmantošana tiks sīkāk aplūkota nākamajā nodaļā par ETL -procesiem.

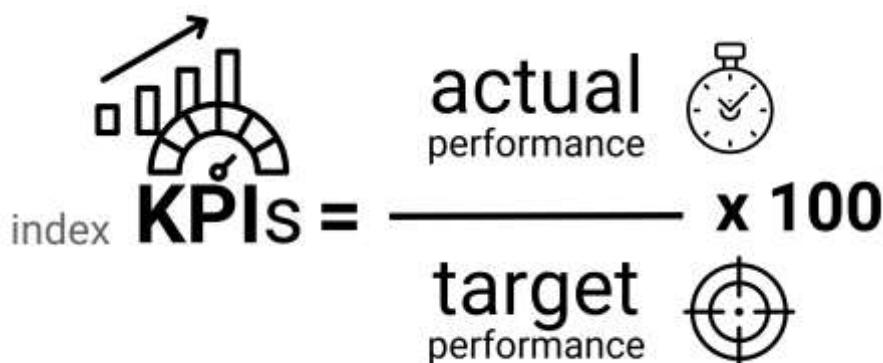
Vizualizācija kļūst par neatņemamu elementu darbā ar datiem būvniecības nozarē - tā palīdz ne tikai "redzēt" datus, bet arī izprast to nozīmi pārvaldības uzdevumu kontekstā. Tomēr, lai vizualizācija būtu patiesi noderīga, ir nepieciešams iepriekš noteikt, kas tieši būtu vizualizējams un kuri rādītāji ir patiešām svarīgi, lai novērtētu projekta izpildi. Šajā gadījumā ir svarīgi izmantot tādus darbības rādītājus kā KPI un ROI. Bez tām pat visskaistākie informācijas paneļi riskē kļūt tikai par "informācijas troksni".

## KPI un ROI

Mūsdienu būvniecības nozarē darbības rādītāju (KPI un ROI) pārvaldībai un to vizualizācijai, izmantojot pārskatus un informācijas paneļus, ir būtiska nozīme produktivitātes un projektu vadības efektivitātes uzlabošanā.

Tāpat kā jebkurā citā uzņēmējdarbībā, arī būvniecībā ir skaidri jādefinē rādītāji, pēc kuriem tiek mērīti panākumi, ieguldījumu atdeve un veikspēja. Iegūstot datus par dažādiem procesiem, uz datiem orientētai organizācijai vispirms ir jāiemācās noteikt galvenos **KPI** (**Key Performance Indicators**) - kvantitatīvos rādītājus, kas atspogulo stratēģisko un darbības mērķu sasniegšanas pakāpi.

Lai aprēķinātu KPI, parasti izmanto formulu ( 7.1-5. attēls), kurā iekļauti faktiskie un plānotie rādītāji. Piemēram, lai aprēķinātu individuālu KPI projektam, darbiniekam vai procesam, faktiskos rādītājus dala ar plānotajiem rādītājiem un rezultātu reizina ar 100%.



**Attēls 7.1-5 KPI tiek izmantoti, lai novērtētu projekta vai procesa panākumus galveno mērķu sasniegšanā.**

Vietnes līmenī var izmantot detalizētākus KPI rādītājus:

- **Galveno atskaites punktu** (pamatu ierīkošana, uzstādīšana, apdare) **laika grafiks** - ļauj kontrolēt atbilstību darba plāniem.
- **Materiālu pārsnieguma procentuālā daļa** - palīdz pārvaldīt iepirkumu un līdz minimumam samazināt izšķērdēšanu.
- **Neplānoto mašīnu dīkstāvju skaits** - ietekmē produktivitāti un izmaksas.

Nepareizu metriku izvēle var novest pie klūdainiem lēmumiem "ko darīt?" ( 2.2.2-5. attēls). Piemēram, ja uzņēmums koncentrējas tikai uz izmaksām uz kvadrātmetru, bet neņem vērā pārbūves izmaksas, materiālu ietaupījumi var novest pie sliktākas kvalitātes un augstākām izmaksām turpmākajos projektos.

Nosakot mērķus, ir svarīgi skaidri noteikt, kas tiek mērīts. Neprecīzs formulējums noved pie nepareiziem secinājumiem un sarežģī kontroli. Aplūkosim veiksmīgu un neveiksmīgu KPI piemērus būvniecībā.

#### **Labi KPI:**

- ☒ "Līdz gada beigām par 10 procentiem samaziniet pārkāsošanas darbu īpatsvaru."
- ☒ "Līdz nākamajam ceturksnim par 15 procentiem palielināt fasāžu montāžas ātrumu,

*nemazinot kvalitāti"*

- ⌚ "Samazināt mašīnu dīkstāves laiku par 20%, optimizējot darba grafikus līdz gada beigām"

Šie rādītāji ir skaidri izmērāmi, tiem ir konkrētas vērtības un termiņi.

#### Slikti KPI:

- ⌚ "Mēs būvēsim ātrāk" (Cik daudz ātrāk? Ko nozīmē "ātrāk"?).
- ⌚ "Mēs uzlabosim konkrēto darbu kvalitāti" (Kā tiesī tiek mērīta kvalitāte?)
- ⌚ "Mēs uzlabosim darbuzņēmēju mijiedarbību būvlaukumā" (Kādi kritēriji liecinās par uzlabojumiem?)

Labs KPI ir tāds, ko var izmērīt un objektīvi novērtēt. Būvniecībā tas ir īpaši svarīgi, jo bez skaidriem rādītājiem nav iespējams uzraudzīt darbības rezultātus un sasniegt stabilus rezultātus.

Papildus KPI ir arī papildu rādītājs ieguldījumu efektivitātes novērtēšanai: **ROI (Return on Investment)** - ieguldījumu atdeves rādītājs, kas atspoguļo peļnas un ieguldīto līdzekļu attiecību. ROI palīdz novērtēt, vai jaunu metožu, tehnoloģiju vai rīku ieviešana ir pamatota: sākot no digitālajiem risinājumiem un automatizācijas (piemēram, 7.3-2. attēls) līdz pat jaunu būvmateriālu izmantošanai. Šis rādītājs palīdz pieņemt pamatotus lēmumus par turpmākiem ieguldījumiem, pamatojoties uz to reālo ietekmi uz uzņēmuma rentabilitāti

Būvniecības projektu vadības kontekstā ROI (ieguldījumu atdevi) var izmantot kā vienu no galvenajiem darbības rādītājiem (KPI), ja uzņēmuma mērķis ir izmērīt projekta, tehnoloģijas vai procesa uzlabošanas ieguldījumu atdevi. Piemēram, ja tiek ieviesta jauna būvniecības pārvaldības metode, ROI var parādīt, cik lielā mērā tā ir uzlabojusi rentabilitāti.

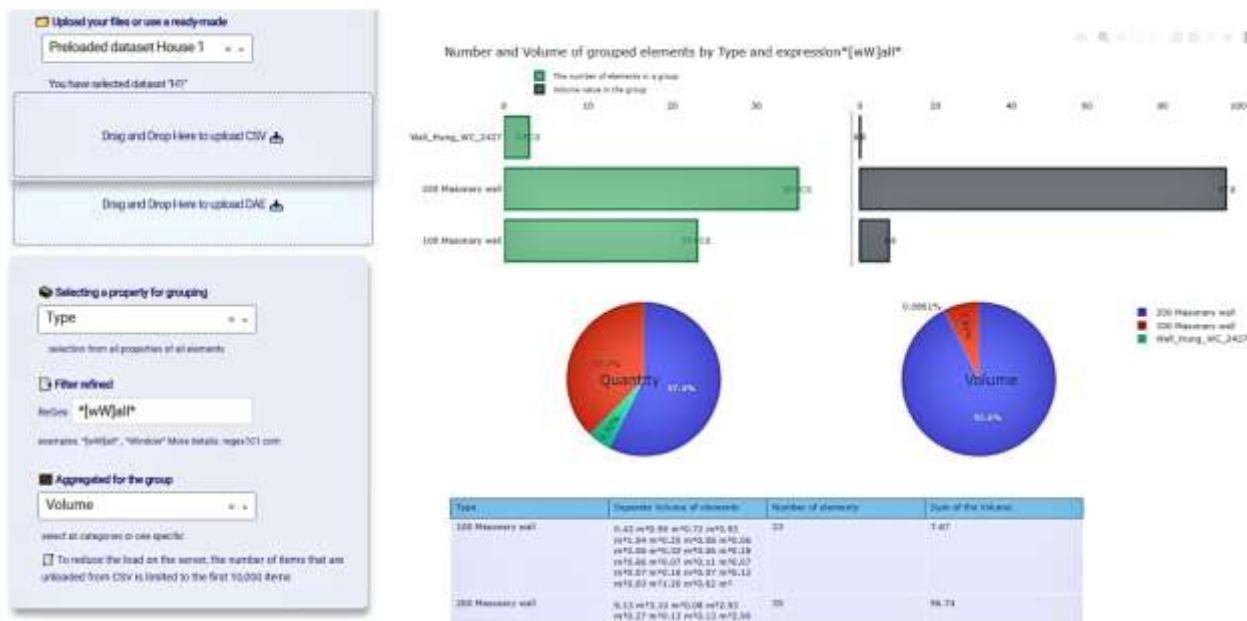
Regulāra KPI un ROI mērīšana, pamatojoties uz datiem, kas iegūti no dažādiem avotiem, piemēram, materiālu patēriņu, darba stundām un izmaksām, ļauj projektu vadībai efektīvi pārvaldīt resursus un ātri pieņemt lēmumus. Šo datu uzglabāšana ilgtermiņā ļauj analizēt nākotnes tendencies un optimizēt procesus.

Lai vizualizētu KPI, ROI un citus rādītājus, tiek izmantotas dažādas diagrammas un grafiki, ko parasti apvieno informācijas paneļos.

### Informācijas paneļi un informācijas paneļi: metriku vizualizēšana efektīvai pārvaldībai

Rādītāju un metriku vizualizēšanai tiek izmantotas dažādas diagrammas un grafiki, kas parasti tiek apvienoti datu vitrīnās un informācijas paneļos. Šie paneļi sniedz centralizētu priekšstatu par projekta vai projekta daļu statusu, parādot galvenos rādītājus (ideālā gadījumā reālajā laikā). Aktualizēti un pastāvīgi atjaunināti paneļi ļauj komandai ātri reaģēt uz izmaiņām.

**Informācijas paneli** ir rīki, kas vizualizē kvantitatīvus novērtējumus, padarot tos viegli pieejamus un saprotamus visiem projekta dalībniekiem.



**Attēls 7.1-6 KPI pārvaldība un to vizualizēšana, izmantojot paneļus, ir galvenais, lai uzlabotu projekta produktivitāti un efektivitāti.**

Šeit ir daži populāru rīku piemēri, ar kuriem varat izveidot paneļus:

- **Power BI** ir Microsoft rīks interaktīvu pārskatu un paneļu izveidei.
  - **Tableau un Google Data Studio** ir jaudīgi rīki datu vizualizēšanai un paneļu izveidei bez nepieciešamības rakstīt kodu.
  - **Plotly** (7.1-6. attēls, 7.2-12. attēls) ir bibliotēka interaktīvu grafiku izveidei, bet Dash ir ietvars tīmekļa lietojumprogrammu izveidei datu analīzei. Tās var izmantot kopā, lai izveidotu interaktīvus paneļus.
  - **Daudzas Python bibliotēkas** ( 7.2-9. attēls - 7.2-11. attēls) - Python ir daudzas atvērtā koda un bezmaksas bibliotēkas datu vizualizācijai, piemēram, Matplotlib, Seaborn, Plotly, Bokeh un citas. Tās var izmantot, lai veidotu grafikus un integrētu tos tīmekļa lietojumprogrammā, izmantojot tādus ietvarus kā Flask vai Django.
  - **JavaScript bibliotēkas:** Jauj izveidot interaktīvus paneļus, izmantojot atvērtā koda JavaScript bibliotēkas, piemēram, D3.js vai Chart.js, un integrēt tos tīmekļa lietojumprogrammā.

Lai novērtētu galvenos rezultatīvos rādītājus un izveidotu informācijas paneļus, jums ir nepieciešami aktuāli dati un skaidrs grafiks informācijas vākšanai un analīzei.

Kopumā KPI, ROI un informācijas paneli būvniecības nozarē veido pamatu analītiskai pieejai projektu

vadībai. Tie ne tikai palīdz uzraudzīt un novērtēt pašreizējo stāvokli, bet arī sniedz vērtīgas atziņas turpmākajiem plānošanas un optimizācijas procesiem - procesiem, kas ir tieši atkarīgi no datu interpretācijas un pareizo un savlaicīgo jautājumu uzdošanas.

## Datu analīze un māksla uzdot jautājumus

Datu interpretācija ir pēdējais analīzes posms, kurā informācija iegūst jēgu un sāk "runāt". Tajā tiek formulētas atbildes uz galvenajiem jautājumiem: "Ko darīt?" un "Kā darīt?" (2.2.2-5. attēls). Šis posms ļauj apkopot rezultātus, identificēt likumsakarības, noteikt cēloņsakarības un izdarīt secinājumus, pamatojoties uz vizualizāciju un statistisko analīzi.

*Iespējams, nav tālu tas laiks, kad tiks saprasts, ka, lai pilnvērtīgi kļūtu par efektīvas kādas no jaunajām lielajām, sarežģītajām pasaules valstīm, kas tagad attīstās, pilsoni, ir tikpat nepieciešams prast rēķināt, domāt vidējo vērtību, maksimumu un minimumu izteiksmē, kā tagad ir nepieciešams prast lasīt un rakstīt [133].*

- Samuels S. Vilkss (Samuel S. Wilkes), citēts 1951. gada prezidenta uzrunā Amerikas Statistikas asociācijai.

Saskaņā ar Apvienotās Karalistes valdības publicēto ziņojumu "Datu analīтика un mākslīgais intelekts valdības projektu īstenošanā" (2024) [83] analītikas datu un mākslīgā intelekta (AI) ieviešana var būtiski uzlabot projektu vadības procesus, palielinot laika un izmaksu prognozēšanas precīzitāti, kā arī samazinot risku un nenoteiktību. Dokumentā uzsvērts, ka publiskās organizācijas, kas izmanto progresīvus analītiskos rīkus, sasniedz augstākus rezultātus infrastruktūras iniciatīvās.

Mūsdieni būvniecības uzņēmējdarbību, kas darbojas ceturtās industriālās revolūcijas apstākļos, kuros valda liela konkurence un zema peļņa, var salīdzināt ar militārajām operācijām. Šeit uzņēmuma izdzīvošana un panākumi ir atkarīgi no resursu un kvalitatīvas informācijas iegūšanas ātruma - un līdz ar to no savlaicīgas un pamatotas lēmumu pieņemšanas (7.1-7. attēls).

Ja datu vizualizācija ir "izlūkošana", kas nodrošina pārskatu, tad datu analīze ir "munīcija", kas nepieciešama rīcībai. Tā sniedz atbildes uz jautājumiem: *ko darīt* un *kā to darīt*, veidojot pamatu konkurences priekšrocību iegūšanai tirgū.

Analītika pārvērš atšķirīgus datus strukturētā un jēgpilnā informācijā, uz kuras pamata var pieņemt lēmumus.

Analītiku un vadītāju uzdevums ir ne tikai interpretēt informāciju, bet arī piedāvāt pamatotus lēmumus, identificēt tendences, noteikt sakarības starp dažādiem datu veidiem un klasificēt tos atbilstoši projekta mērķiem un specifikai. Izmantojot vizualizācijas rīkus un statistiskās analīzes metodes, viņi pārvērš datus par uzņēmuma stratēģisku vērtību.



**Attēls 7.1-7 Tā ir datu analīze, kas galu galā pārvērš savākto informāciju par avotu lēmumu pieņemšanai.**

Lai pieņemtu patiesi pamatotus lēmumus analīzes procesā, ir jāiemācās pareizi formulēt jautājumus, kas tiek uzdoti par datiem. Šo jautājumu kvalitāte tieši ietekmē iegūto atziņu dzījumu un līdz ar to arī vadības lēmumu kvalitāti.

*Pagātne eksistē tikai tiktāl, ciktāl tā ir klātesoša šodienas ierakstos. Un to, ko šie ieraksti atspoguļo, nosaka mūsu uzdotie jautājumi. Nav citas vēstures kā tikai šī [134].*

- Džons Arčibalds Vīlers, fiziķis 1982

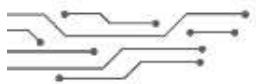
Māksla uzdot padziļinātus jautājumus un kritiski domāt ir būtiska prasme darbā ar datiem. Lielākā daļa cilvēku mēdz uzdot vienkāršus, virspusējus jautājumus, uz kuriem atbildēt prasa maz pūlu. Tomēr patiesa analīze sākas ar jēgpilniem un pārdomātiem jautājumiem, kas var atklāt slēptās sakarības un cēloņsakarības informācijā, kura var būt paslēpta aiz vairākiem argumentācijas slāņiem.

Saskaņā ar pētījumu "Data-Driven Transformation: Accelerating at Scale Now" (BCG, 2017) [135], veiksmīgai digitālajai transformācijai ir nepieciešami ieguldījumi analītikas spējās, pārmaiņu vadības programmās un biznesa mērķu saskaņošana ar IT iniciatīvām. Uzņēmumiem, kas veido uz datiem balstītu kultūru, ir jāiegulda līdzekļi datu analīzes spējās un jāuzsāk pārmaiņu pārvaldības programmas, lai iedzīvinātu jaunu domāšanu, uzvedību un darba metodes.

Neieguldīt līdzekļus analītiskās kultūras attīstībā, datu rīku uzlabošanā un speciālistu apmācībā, uzņēmumi turpinās riskēt pieņemt lēmumus, pamatojoties uz novecojušu vai nepilnīgu informāciju, vai palauties uz HiPPO vadītāju subjektīviem viedokļiem ( 2.1-9. attēls).

Apzinoties analītikas un informācijas paneļu nozīmīgumu un nepieciešamību tos pastāvīgi atjaunināt, vadība neizbēgami saprot, cik svarīgi ir automatizēt analītiskos procesus. Automatizācija palielina lēmumu pieņemšanas ātrumu, samazina cilvēciskā faktora ietekmi un nodrošina datu atbilstību. Eksponenciāli pieaugot informācijas apjomiem, ātrums kļūst ne tikai par konkurences priekšrocību, bet arī par galveno ilgtspējīgu panākumu faktoru.

Datu analīzes un apstrādes procesu automatizācija kopumā ir nesaraujami saistīta ar ETL tēmu (Extract, Transform, Load). Tāpat kā automatizācijas procesā ir nepieciešams pārveidot datus, arī ETL procesā dati tiek iegūti no dažādiem avotiem, pārveidoti atbilstoši nepieciešamajām prasībām un ielādēti mērķa sistēmās tālākai izmantošanai.



## NODAĻA 7.2.

### DATU PLŪSMA BEZ MANUĀLAS PIEPŪLES: KĀPĒC NEPIECIEŠAMS ETL

#### ETL automatizācija: izmaksu samazināšana un datu apstrādes paātrināšana

Kad galvenie darbības rādītāji (KPI) vairs nepalielinās, neraugoties uz datu apjoma un komandas lieluma pieaugumu, uzņēmuma vadība neizbēgami sāk apzināties nepieciešamību automatizēt procesus. Agrāk vai vēlāk šī apziņa kļūst par stimulu uzsākt kompleksu automatizāciju, kuras galvenais mērķis ir samazināt procesu sarežģītību, paātrināt apstrādi un samazināt atkarību no cilvēciskā faktora.

Saskaņā ar McKinsey pētījumu "Kā izveidot datu arhitektūru, lai veicinātu inovācijas - šodien un rīt" (2022) [136], uzņēmumi, kas izmanto datu plūsmas arhitektūru, iegūst ievērojamas priekšrocības, jo tie var analizēt informāciju reālajā laikā. Straumēšanas tehnoloģijas ļauj tieši analizēt reāllaika ziņojumus un piemērot prognozējošo apkopi ražošanā, analizējot reāllaika sensoru datus.

Procesu vienkāršošana ir automatizācija, kad tradicionālās manuālās funkcijas tiek aizstātas ar algoritmiem un sistēmām.

Jautājums par automatizāciju jeb, pareizāk sakot, "cilvēka lomas samazināšanu datu apstrādē" ir neatgriezenisks un ļoti jutīgs process katrā uzņēmumā. Jebkuras profesionālās jomas speciālisti bieži vien vilcinās pilnībā atklāt savas darba metodes un smalkumus kolēģiem optimizētājiem, apzinoties risku zaudēt darbu strauji augošajā tehnoloģiskajā vidē.

*Ja vēlaties iegūt ienaidniekus, mēģiniet mainīt lietas [137].*

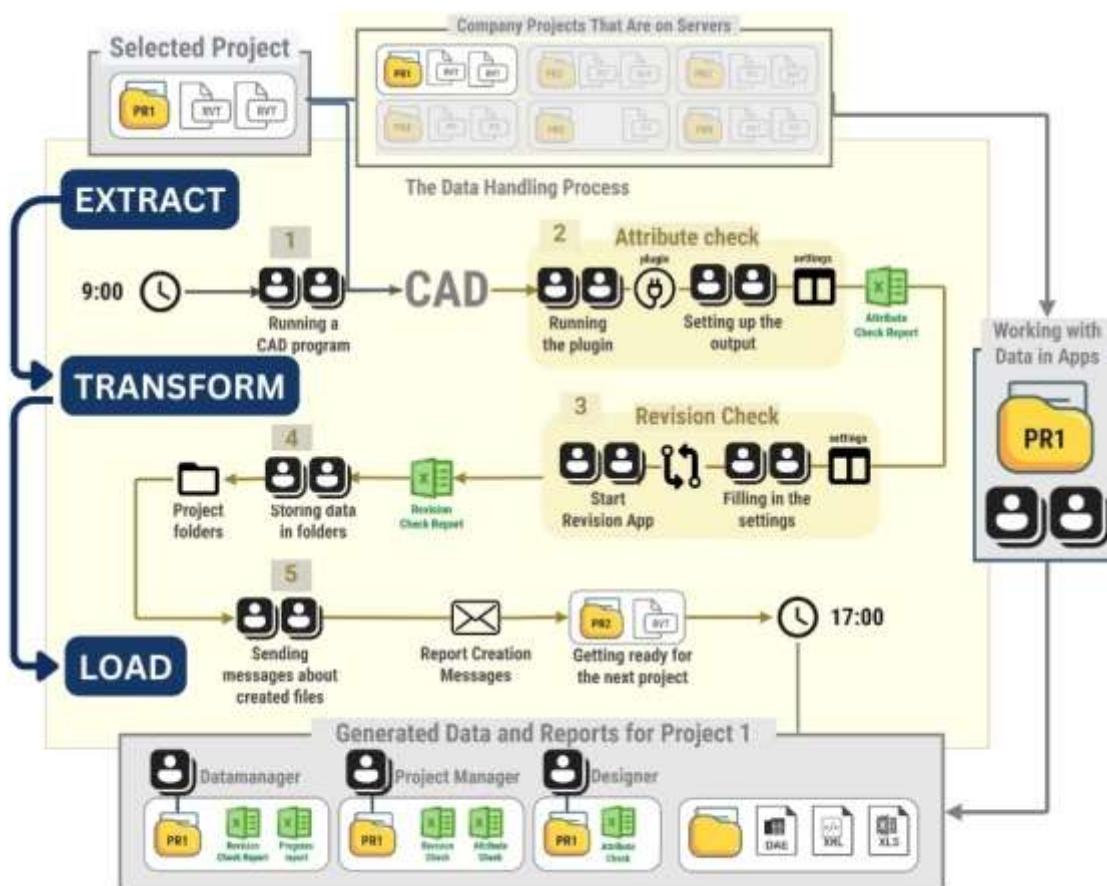
- Vudro Vilsons, runa tirgotāju kongresā, Detroita, 1916. g.

Neraugoties uz acīmredzamajām automatizācijas priekšrocībām, daudzu uzņēmumu ikdienas praksē joprojām ir liels roku darba īpatsvars, jo īpaši inženiertehnisko datu jomā. Lai ilustrētu pašreizējo situāciju, aplūkosim tipisku secīgas datu apstrādes piemēru šādos procesos.

Manuālu datu apstrādi var ilustrēt ar piemēru par mijiedarbību ar informāciju, kas iegūta no CAD datubāzēm. Tradicionālā datu apstrāde ("manuālais" ETL -process) CAD (BIM) nodalās atribūtu tabulu izveidei vai dokumentācijas izveidei, pamatojoties uz projektēšanas datiem, notiek šādā secībā (7.2-1. attēls):

1. Manuāla **ieguve (Extract)**: lietotājs manuāli atver projektu, iedarbinot CAD lietojumprogrammu (BIM) (1. attēls 7.2-1. solis).
2. **Verifikācija**: nākamajā posmā parasti manuāli tiek palaisti vairāki spraudņi vai palīgprogrammas, lai sagatavotu datus un novērtētu to kvalitāti (7.2-1. attēls, 2-3. solis).

3. Manuāla **transformācija (Transform)**: pēc sagatavošanas sākas datu apstrāde, kas prasa manuālu darbību ar dažādiem programmatūras rīkiem, kuros dati tiek sagatavoti augšpielādei (7.2-1. attēls, 4. solis).
4. Manuāla **augšpielāde (Load)**: manuāla konvertēto datu augšpielāde uz ārējām sistēmām, datu formātiem un dokumentiem (7.2-1. attēls, 5. solis).



Attēls 7.2-1 Tradicionālo manuālo ETL apstrādi ierobežo individuālā tehnika vēlmes un fiziskās iespējas.

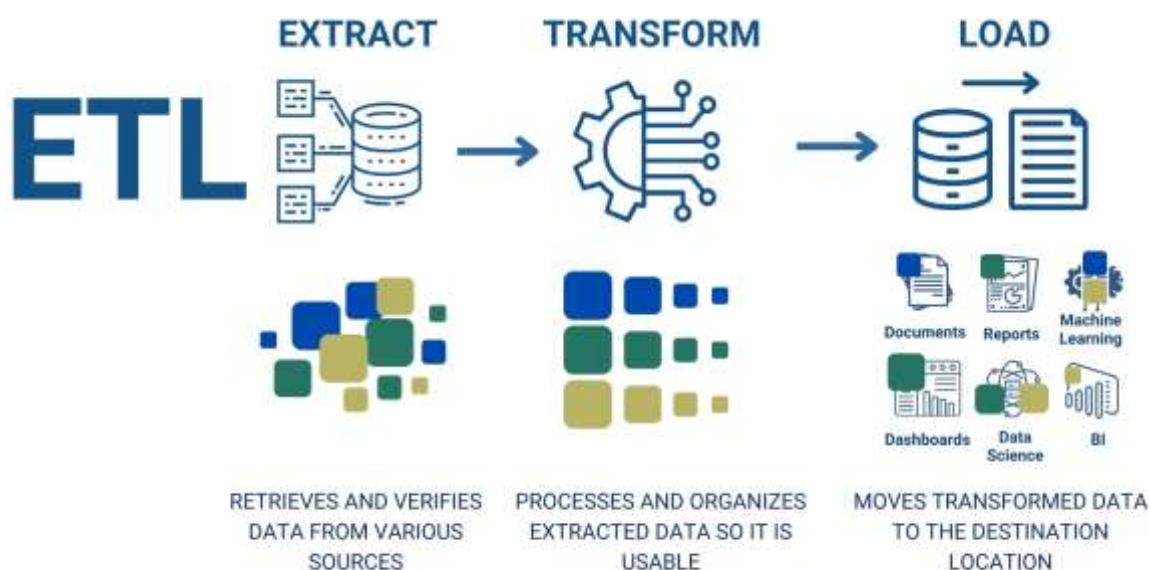
Šāda darbplūsma ir klasiska ETL -procesa - ekstrakcijas, transformācijas un ielādes (ETL) - piemērs. Atšķirībā no citām nozarēm, kurās automātiskie ETL cauruļvadi jau sen ir kļuvuši par standartu, būvniecības nozarē joprojām dominē roku darbs, kas palēnina procesus un palielina izmaksas

**ETL (Extract, Transform, Load)** ir datu ieguves process no dažādiem avotiem, to pārveidošana vēlamajā formātā un ielāde mērķa sistēmā turpmākai analīzei un izmantošanai.

**ETL** ir process, kas apzīmē trīs galvenos datu apstrādes komponentus: izvilkšana, pārveidošana un ielāde (7.2-2. attēls):

- **Izraksts** - datu izraksts no dažādiem avotiem (failiem, datubāzēm, API).
- **Pārveidot** - datu tīrišana, apkopošana, normalizēšana un loģiskā apstrāde.
- **Ielādējet** - ielādējet strukturētu informāciju datu krātuvē, pārskatā vai BI sistēmā.

Iepriekš grāmatā ETL koncepcija tika skarta tikai dažkārt: nestrukturēta dokumenta pārveidošanā strukturētā tabulas formātā (4.1-1. attēls), prasību formalizēšanas kontekstā, lai sistematizētu gan dzīves, gan biznesa procesu uztveri (4.4-20. attēls), kā arī datu validēšanas un datu apstrādes automatizācijā no CAD risinājumiem. Tagad aplūkosim ETL sīkāk tipisku darba plūsmu kontekstā.



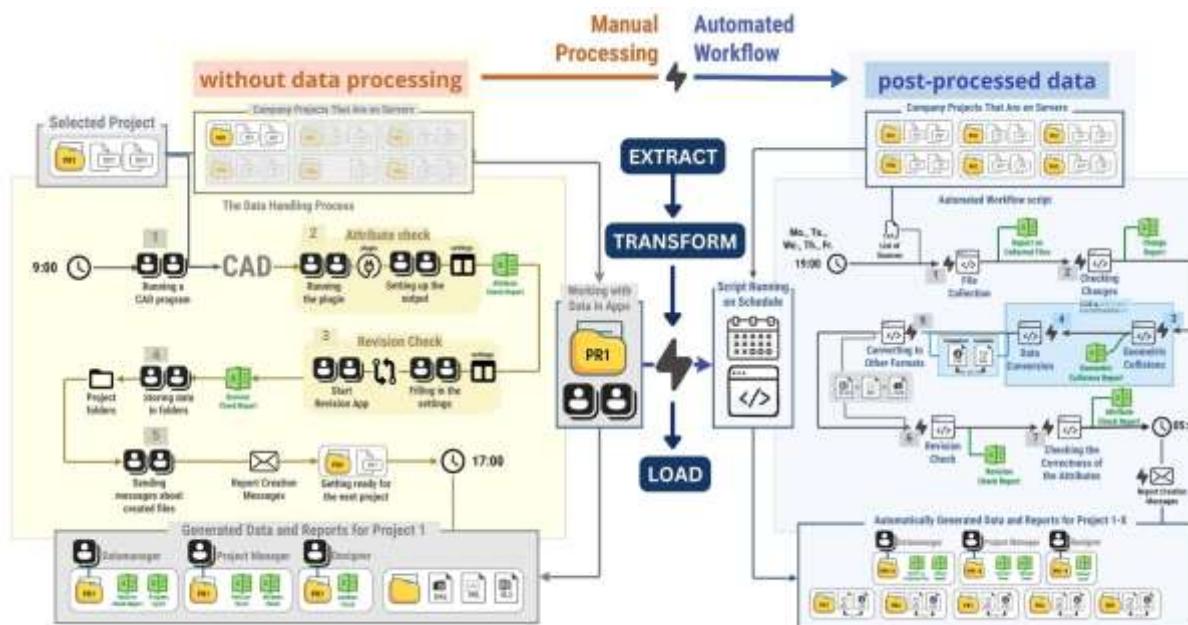
Attēls 7.2-2 ETL automatizē atkārtotus datu apstrādes uzdevumus.

Manuāls vai daļēji automatizēts ETL - process nozīmē, ka vadītājs vai tehniskais darbinieks visus posmus - no datu vākšanas līdz atskaņai - pārvalda manuāli. Šāds process aizņem daudz laika, jo īpaši, ja darba laiks ir ierobežots (piemēram, no 9:00 līdz 17:00).

Uzņēmumi bieži cenšas atrisināt zemas efektivitātes un lēna ātruma problēmu, iegādājoties modulārus integrētus risinājumus (ERP, PMIS, CPM, CAFM u. c.), kurus pēc tam tālāk attīsta ārējie piegādātāji un konsultanti. Taču šie pārdevēji un trešo pušu izstrādātāji bieži vien klūst par kritisku atkarības punktu: to tehniskie ierobežojumi tieši ietekmē visas sistēmas un uzņēmuma efektivitāti kopumā, kā sīkāk aprakstīts iepriekšējās nodajās par patentētām sistēmām un formātiem. Sadrumstalotības un atkarības radītās problēmas tika detalizēti aplūkotas nodajā "Kā būvniecības uzņēmumi slīgst datu haosā".

Ja uzņēmums nav gatavs ieviest lielu modulāru platformu no kāda piegādātāja, tas sāk meklēt alternatīvus automatizācijas veidus. Viens no tiem ir izstrādāt savu moduļu atvērto ETL -konveijeri, kur katrs posms (ieguve, transformācija, validācija, ielāde) tiek īstenošs kā skripti, kas tiek izpildīti pēc grafika.

Tās pašas ETL darba plūsmas automatizētajā versijā ( 7.2-1. attēls) darba process izskatās kā modulārs kods, kas sākas ar datu apstrādi un to pārveidošanu atvērtā strukturētā formā. Kad strukturētie dati ir saņemti, automātiski pēc grafika tiek palaisti dažādi skripti vai moduļi, lai pārbaudītu izmaiņas, transformētu un nosūtītu ziņojumus ( 7.2-3. attēls).



Attēlā 7.2-3 Kreisajā pusē ir manuālā apstrāde, bet labajā - automātiskais process, kuru atšķirībā no tradicionālās manuālās apstrādes neierobežo lietotāja iespējas.

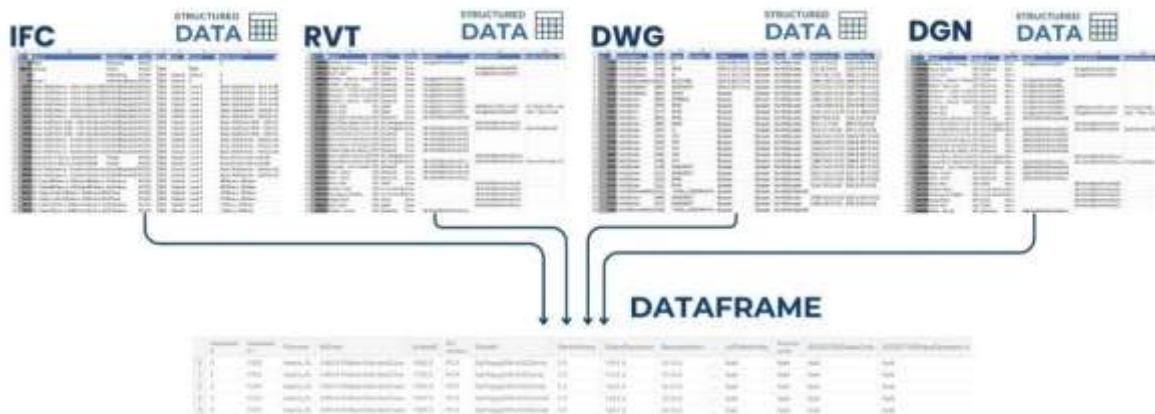
Automatizētā darba plūsmā datu apstrādi vienkāršo ET(L) datu pirmapstrāde: strukturēšana un apvienošana.

Tradicionālajās apstrādes metodēs speciālisti strādā ar datiem "tādiem, kādi tie ir", t.i., tādiem, kādi tie ir iegūti no sistēmām vai programmatūras. Turpretī automatizētajos procesos dati bieži vien vispirms iziet cauri ETL -payplane, kur tie tiek sagatavoti konsekventā struktūrā un formātā, kas piemērots tālākai izmantošanai un analīzei.

Aplūkosim praktisku ETL piemēru, demonstrējot datu tabulas validēšanas procesu, kas aprakstīts nodajā "datu validēšana un validēšanas rezultāti" (4.4-13. attēls). Šim nolūkam mēs izmantojam Pandas bibliotēku kopā ar LLM automatizētajiem datu analīzes un apstrādes procesiem.

## ETL Izraksts: datu vākšana

Pirmais ETL procesa posms - ekstrakts) - sākas ar koda rakstīšanu, lai apkopotu datu kopas, kas tiks tālāk pārbaudītas un apstrādātas. Lai to izdarītu, mēs skenēsim visas produkcijas servera mapes, savāksim noteikta formāta un saturu dokumentus un pēc tam tos pārveidosim strukturētā formā. Šis process ir detalizēti aplūkots nodaļās "Nestrukturētu un teksta datu pārvēršana strukturētā formā" un "CAD datu (BIM) pārvēršana strukturētā formā" ( 4.1-1. attēls - 4.1-12. attēls).



Attēls 7.2-4 Konvertējet CAD datus (BIM) vienā lielā datu rāmī, kas satur visas projekta dajas.

Kā ilustratīvu piemēru mēs izmantojam datu ielādes soli Extract un iegūstam visu CAD- (BIM-) projektu tabulu (7.2-4. attēls) ( 7.2-4. attēls), izmantojot apgrieztās inženierijas atbalstītu konvertētāju [138] RVT un IFC formātiem, lai iegūtu strukturētas tabulas no visiem projektiem un apvienotu tās vienā lielā DataFrame tabulā.

```

1 import os
2 import subprocess
3 import time
4 import pandas as pd
5
6 path_conv = r'C:\DDC_2023\\' # Where RvtExporter.exe\IfcExporter.exe is located
7 path = r'C:\IFCprojects\\' # Where Revit\IFC files are stored
8
9 def convert_and_wait(path_conv, exporter_name, file_path, extension):
10     # Start the conversion process
11     subprocess.Popen([os.path.join(path_conv, exporter_name),
12                      file_path], cwd=path_conv)
13     output_file = os.path.join(path,
14                               f"{os.path.splitext(os.path.basename(file_path))[0]}_{extension}.xlsx")
15
16     # Conversion process for RVT and IFC files
17     for file in os.listdir(path):
18         full_path = os.path.join(path, file)
19         if file.endswith('.ifc'):
20             convert_and_wait(path_conv, 'IfcExporter.exe', full_path, 'ifc')
21         elif file.endswith('.rvt'):
22             convert_and_wait(path_conv, 'RvtExporter.exe', full_path, 'rvt')
23
24     # Combine converted Excel files into one dataframe
25 df = pd.concat([pd.read_excel(os.path.join(path, f)) for f in os.listdir(path)
26 if f.endswith('.xlsx')], ignore_index=True)

```

**Attēls 7.2-5 Konvertēšana, izmantojot Python kodu un SDK apgrieztās inženierijas rīku RVT un IFC failus vienā lielā strukturētā (df) DataFrame.**

Pandas DataFrame var ielādēt datus no dažādiem avotiem, tostarp CSV teksta failiem, Excel izklājlapām, JSON - un XML - failiem, lieliem datu glabāšanas formātiem, piemēram, Parquet un HDF5, un no MySQL, PostgreSQL, SQLite, Microsoft SQL Server, Oracle un citām datubāzēm. Turklāt Pandas atbalsta datu ielādi no API, tīmekļa lapām, mākoņpakaļpojumiem un glabāšanas sistēmām, piemēram, Google BigQuery, Amazon Redshift un Snowflake.

- ❷ Lai uzrakstītu kodu, ar ko izveidot savienojumu un vākt informāciju no datubāzēm, nosūtiet līdzīgu teksta pieprasījumu uz LLM tērzēšanas istabu (CHATGP, LLaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkuru citu):

Lūdzu, uzrakstiet piemēru, kā izveidot savienojumu ar MySQL un konvertēt datus uz.



LLM atbilde :

```

1 import pandas as pd
2 import mysql.connector
3
4 # Using a context manager to ensure
5 # that the connection is closed automatically
6 with mysql.connector.connect(host='hostname', database='dbname',
7 user='username', password='password') as conn:
8     df = pd.read_sql('SELECT * FROM tablename LIMIT 10', conn)

```

Attēls 7.2-6 Piemērs, kā, izmantojot Python , izveidot savienojumu ar MySQL datubāzi un importēt datus no MySQL datubāzes uz DataFrame .

legūtais kods ( 7.2-5. attēls, 7. attēls).2-6) var palaist vienā no populārākajām IDE (integrētajām izstrādes vidēm), kurās minējām iepriekš, bezsaistes režīmā: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populāros tiešsaistes rīkus: Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

Ievietojot daudzformāta datus mainīgajā "df" ( 7.2-5. attēls - 25. rinda; 7.2-6. attēls - 8. rinda), mēs konvertējām datus Pandas DataFrame formātā, kas ir viena no populārākajām datu apstrādes struktūrām un kas ir divdimensiju tabula ar rindām un kolonnām. Par citiem datu glabāšanas formātiem, ko izmanto ETL -Pipeline, piemēram, Parquet, Apache ORC, JSON, Feather, HDF5, kā arī par modernām datu noliktavām mēs vairāk runāsim nodaļā "Datu glabāšana un pārvaldība būvniecības nozarē" ( 8.1-2. attēls).

Pēc datu ieguves un strukturēšanas posma (Extract) tiek izveidots vienots informācijas masīvs ( 7.2-5. attēls, 7.2-6. attēls), kas ir gatavs turpmākai apstrādei. Tomēr pirms šo datu ielādēšanas mērķa sistēmās vai izmantošanas analīzei ir jānodrošina to kvalitāte, integritāte un atbilstība noteiktajām prasībām. Tieši šajā posmā notiek datu transformācija (Transform), kas ir galvenais solis, lai nodrošinātu turpmāko secinājumu un lēmumu ticamību.

## ETL Transform: validācijas un transformācijas noteikumu piemērošana.

Pārveidošanas posmā dati tiek apstrādāti un pārveidoti. Šis process var ietvert pareizības pārbaudi, normalizāciju, trūkstošo vērtību aizpildīšanu un validāciju, izmantojot automatizētus rīkus

Saskaņā ar PwC pētījumu "Datu vadīti. Kas studentiem nepieciešams, lai gūtu panākumus strauji mainīgajā uzņēmējdarbības pasaule" (2015) [9], mūsdienu revīzijas uzņēmumi atsakās no nejaušas datu pārbaudes un pāriet uz milzīgu informācijas apjomu analīzi, izmantojot automatizētus rīkus. Šāda pieeja ļauj ne tikai identificēt neatbilstības pārskatos, bet arī piedāvāt ieteikumus uzņēmējdarbības procesu optimizācijai.

Būvniecībā līdzīgas metodes var izmantot, piemēram, projektēšanas datu automātiskai validācijai, būvniecības kvalitātes kontrolei un būvuzņēmēju darbības novērtēšanai. Viens no rīkiem datu apstrādes automatizēšanai un paātrināšanai ir regulāru izteiksmju izmantošana (RegEx) datu transformācijas fāzē (Transform) ETL procesā. RegEx ļauj efektīvi pārbaudīt datu virknes, atklāt neatbilstības un nodrošināt informācijas integritāti ar minimālu resursu patēriņu. Sīkāka informācija par RegEx (4.4-7. attēls) tika aplūkota nodaļā "Prasību tulkošana strukturētā formā".

Aplūkosim praktisku piemēru: īpašumu objektu pārvaldības sistēmā (RPM) pārvaldnieks nosaka prasības objektu galvenajiem atribūtiem (7.2-7. attēls). Transformācijas posmā ir jāapstiprina šādi parametri:

- objektu identifikatoru formātu pārbaude (atribūts "ID")
- aizvietošanas garantijas perioda vērtību kontrole (atribūts "Garantijas periods").
- elementa nomaiņas cikla pārbaude (atribūts "Tehniskās apkopes prasības").

Property Manager: Long-term Management				
ID	Element	Warranty Period	Replacement Cycle	Maintenance Requirements
W-NEW	Window	-	20 years	Annual Inspection
W-OLD1	Window	8 years	15 years	Biannual Inspection
W-OLD2	Window	8 years	15 years	Biannual Inspection
D-122	Door	15 years	25 years	Biennial Varnishing

Attēls 7.2-7 Verifikācija kvalitāte sākas ar prasību noteikšanu atribūtiem un to robežvērtībām.

Piemēram, lai iestatītu parametru validācijas robežvērtības, pieņemsim, ka no pieredzes zinām, ka atribūta "ID" derīgās vērtības var būt tikai virknes vērtības "W-NEW", "W-OLD1" vai "D-122" vai līdzīgas vērtības, kur pirmā rakstzīme ir burts, kam seko domuzīme un pēc tam trīs alfabēta zīmes "NEW", "OLD" vai jebkurš trīsciparu skaitlis (7.2-7. attēls). Šo identifikatoru validēšanai var izmantot šādu regulāro izteiksmi (RegEx):

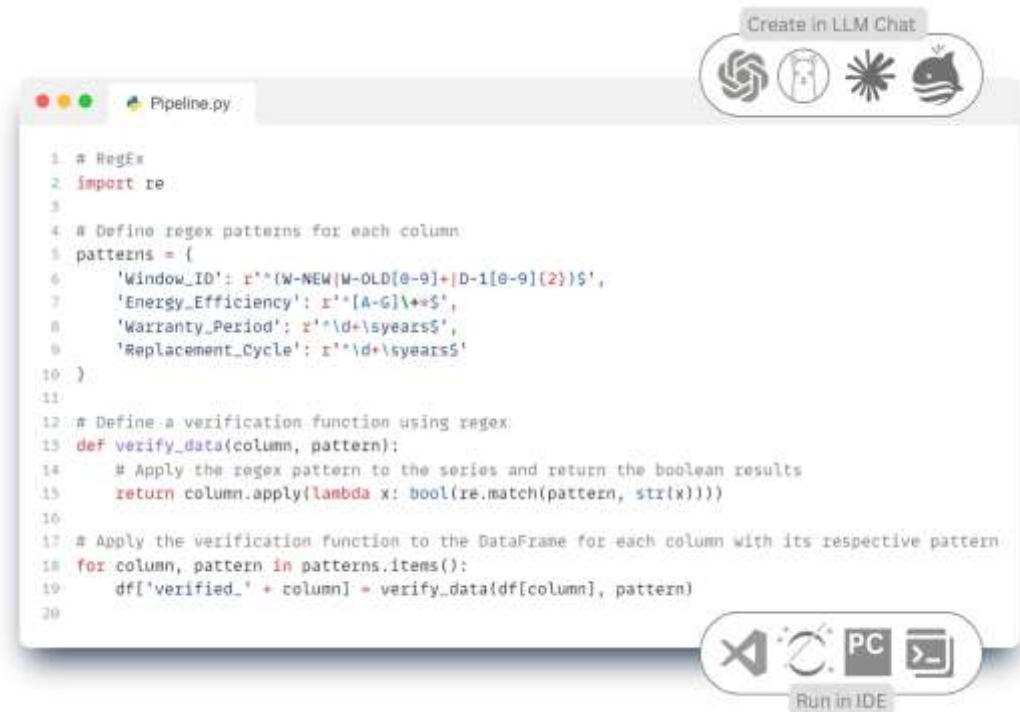
```
^W-NEW$ | ^W-OLD[0-9]+$ | ^D-1[0-9]{2}$
```

Šī veidne nodrošina, ka visi identifikatori datos atbilst norādītajiem kritērijiem. Ja kāda vērtība neiztur pārbaudi, sistēma reģistrē kļūdu. Lai izveidotu Python -kodu datu pārveidošanai un izmantotu iegūtos datus rezultātu tabulas izveidei, vienkārši formulējiet vaicājumu LLM tērzēšanā.

#### 💡 Teksta pieprasījums LLM:

Uzrakstiet kodu, lai validētu datu rāmja kolonas, izmantojot regulārās izteiksmes, kas pārbauda identifikatorus formātā "W-NEW" vai "W-OLD", izmantojot RegEx, energoefektivitāti ar burtiem no "A" līdz "G", garantijas periodu un nomaiņas ciklu ar skaitliskām vērtībām gados. ↴

#### 💡 LLM atbilde :



```
1 # Regex
2 import re
3
4 # Define regex patterns for each column
5 patterns = {
6     'Window_ID': r'^(W-NEW|W-OLD[0-9]+|D-1[0-9]{2})$',
7     'Energy_Efficiency': r'^[A-G]\w*$',
8     'Warranty_Period': r'^\d+\syears$',
9     '#replacement_Cycle': r'^\d+\syears$'
10 }
11
12 # Define a verification function using regex
13 def verify_data(column, pattern):
14     # Apply the regex pattern to the series and return the boolean results
15     return column.apply(lambda x: bool(re.match(pattern, str(x))))
16
17 # Apply the verification function to the DataFrame for each column with its respective pattern
18 for column, pattern in patterns.items():
19     df['verified_' + column] = verify_data(df[column], pattern)
```

Attēls 7.2-8 Kods automatizē validēšanas procesu, piemērojot RegEx veidnes datu rāmja parametru kolonnām .

Iepriekš minētajā automātiski ģenerētajā Python -kodā ( 7.2-8. attēls) tiek izmantota "re" bibliotēka (RegEx regulārās izteiksmes), lai definētu funkciju, kas pārbauda katru datu elementa atribūtu DataFrame. Katrai norādītajai kolonai (atribūtam) funkcija piemēro RegEx modeli, lai pārbaudītu, vai katrs ieraksts atbilst gaidītajam formātam, un pievieno rezultātus kā jaunas vērtības (False/True) jaunajā DataFrame atribūta kolonā.

Šāda automatizēta validācija nodrošina formālu datu atbilstību, un to var izmantot kā daļu no kvalitātes kontroles sistēmas transformācijas posmā.

Pēc veiksmīgas transformēšanas posma un kvalitātes pārbaudes pabeigšanas dati ir gatavi augšupielādei mērķa sistēmās. Pārveidotos un apstiprinātos datus var augšupielādēt CSV, JSON, Excel, datubāzēs un citos formātos turpmākai izmantošanai. Atkarībā no uzdevuma rezultātus var arī attēlot ziņojumos, grafikos vai analītiskos paneļos.

## ETL Load: Vizualizējet rezultātus diagrammās un grafikos.

Pēc Transform posma pabeigšanas, kad dati ir strukturēti un pārbaudīti, pēdējais posms ir Load, kurā datus var gan ielādēt mērķa sistēmā, gan vizualizēt analīzei. Datu vizuālais attēlojums ļauj ātri identificēt novirzes, analizēt sadalījumu un paziņot galvenos secinājumus visiem projekta dalībniekiem, arī tiem, kam nav tehniskas sagatavotības.

Tā vietā, lai informāciju sniegtu tabulu un skaitļu veidā, mēs varam izmantot infografikas, grafikus un informācijas paneļus (dashboards). Viens no izplatītākajiem un elastīgākajiem rīkiem strukturētu datu vizualizēšanai Python ir Matplotlib bibliotēka ( 7.2-9. attēls, 7.2-10. attēls). Tā ļauj veidot statiskas, animētas un interaktīvas diagrammas, kā arī atbalsta plašu diagrammu tipu klāstu.

- Lai vizualizētu RPM sistēmas atribūtu pārbaudes rezultātus ( 7.2-7. attēls), varat izmantot šādu valodas modeļa pieprasījumu:

Uzrakstiet kodu, lai vizualizētu DataFrame datus, iepriekš ( 7.2-7. attēls), ar histogrammu rezultātiem, lai parādītu kļūdu biežumu atribūtā. ↵

- ❷ LLM atbilde koda formā un gatavu vizualizāciju tieši LLM tērzēšanas telpā par koda izpildes rezultātiem :

The screenshot shows a software interface for generating code and visualizations. At the top, there's a toolbar with icons for 'Create in LLM Chat' and four decorative icons. Below that is a code editor window titled 'Pipeline.py' containing the following Python code:

```

1 # Re-importing necessary libraries for visualization
2 import matplotlib.pyplot as plt
3 import seaborn as sns
4
5 # Visualization 1: Bar Chart
6 plt.figure(figsize=(10, 6))
7 df_visual.plot(kind='bar', stacked=True, color=['green', 'red'])
8 plt.title('Data Verification Summary - Bar Chart')
9 plt.xlabel('Data Categories')
10 plt.ylabel('Count')
11 plt.xticks(rotation=45)
12 plt.tight_layout()
13 plt.show()

```

Below the code editor is a 'Run in IDE' button with icons for a terminal, a circular progress bar, a PC monitor, and a file. The main area displays the generated visualization, which is a stacked bar chart titled 'Data Verification Summary - Bar Chart'. The Y-axis is labeled 'Count (of values checked)' and ranges from 0 to 5. The X-axis is labeled 'Data Categories' and lists four categories: 'Window\_ID', 'Energy\_Efficiency', 'Warranty\_Period', and 'Replacement\_Cycle'. Each category has a stacked bar consisting of two segments: a green segment at the bottom representing 'Passed' values and a red segment at the top representing 'Failed' values. The legend indicates that green represents 'Passed' and red represents 'Failed'.

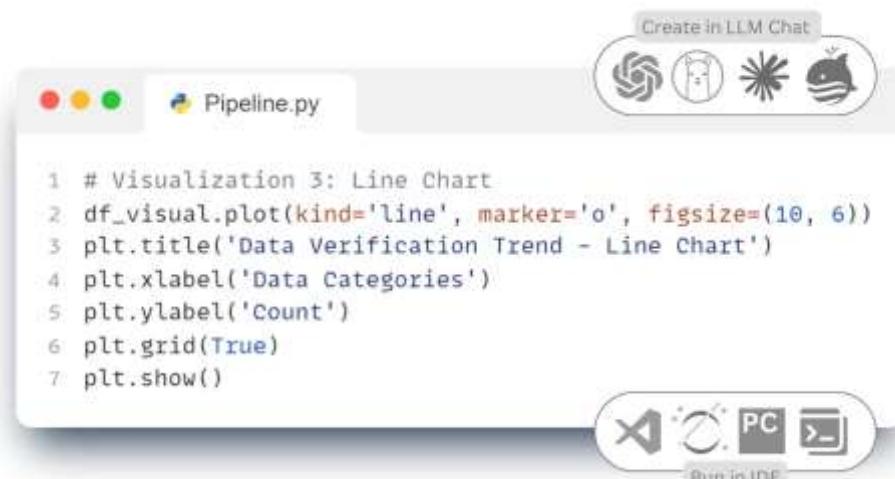
Data Category	Passed (Green)	Failed (Red)	Total Count
Window_ID	4	1	5
Energy_Efficiency	3	2	5
Warranty_Period	4	1	5
Replacement_Cycle	4	1	5

Attēls 7.2-9. Transformēšanas posma rezultātu vizualizācija , pārbaudot RPM sistēmas atribūtu vērtības (attēls 7.2-7) kā histogrammu Load posmā .

- Pastāv daudzas atvērtā koda un bezmaksas vizualizācijas bibliotēkas, kas ļauj attēlot strukturētus datus dažādos formātos. Turpināsim vizualizēt rezultātus ar cita veida grafiku, izmantojot šādu uzvedni tērzēšanas vietnē:

Uzzīmējiet tos pašus datus kā līniju grafiku ↪

- LLM atbilde :



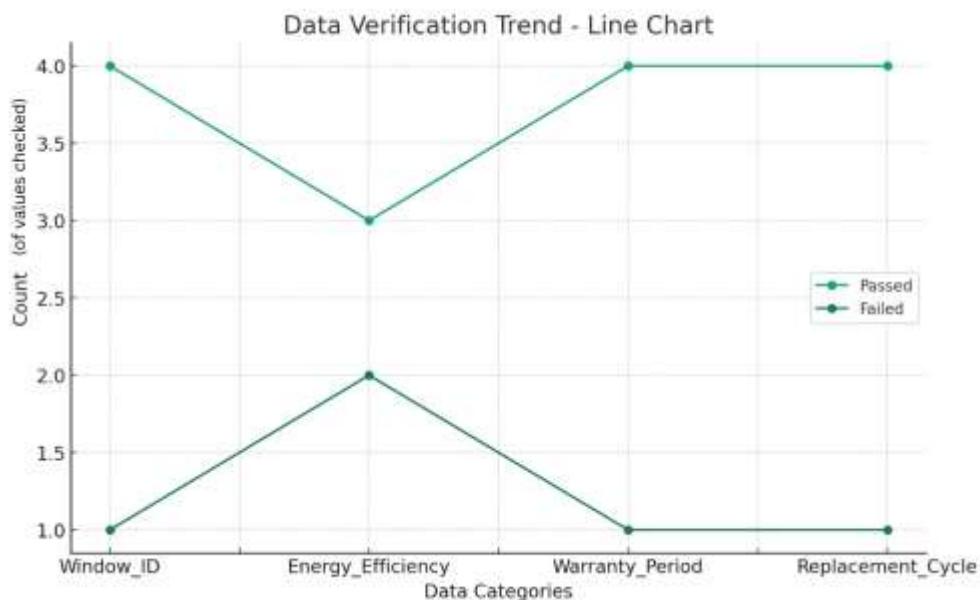
The screenshot shows a user interface for generating code and visualizations. At the top, there's a toolbar with icons for file operations and a "Create in LLM Chat" button. Below it, a "Pipeline.py" tab is open, displaying the following Python code:

```

1 # Visualization 3: Line Chart
2 df_visual.plot(kind='line', marker='o', figsize=(10, 6))
3 plt.title('Data Verification Trend - Line Chart')
4 plt.xlabel('Data Categories')
5 plt.ylabel('Count')
6 plt.grid(True)
7 plt.show()

```

At the bottom right, there are icons for running the code in an IDE and a "Run in IDE" button.



**Attēls 7.2-10 Vizualizācija validācijas datiem (attēls 7.2-8) kā līnijdiagramma, kas iegūta, izmantojot Matplotlib bibliotēku .**

Ir pieejamas daudzas atvērtā koda un bezmaksas vizualizācijas bibliotēkas, piemēram:

- Seaborn - statistikas grafikiem ( 7.2-11. attēls).
- Plotly - interaktīvām tīmekļa vizualizācijām ( 7.2-12. attēls, 7.1-6. attēls).
- Altair - deklaratīvai vizualizācijai
- Dash vai Streamlit - lai izveidotu pilnīgus paneļus.

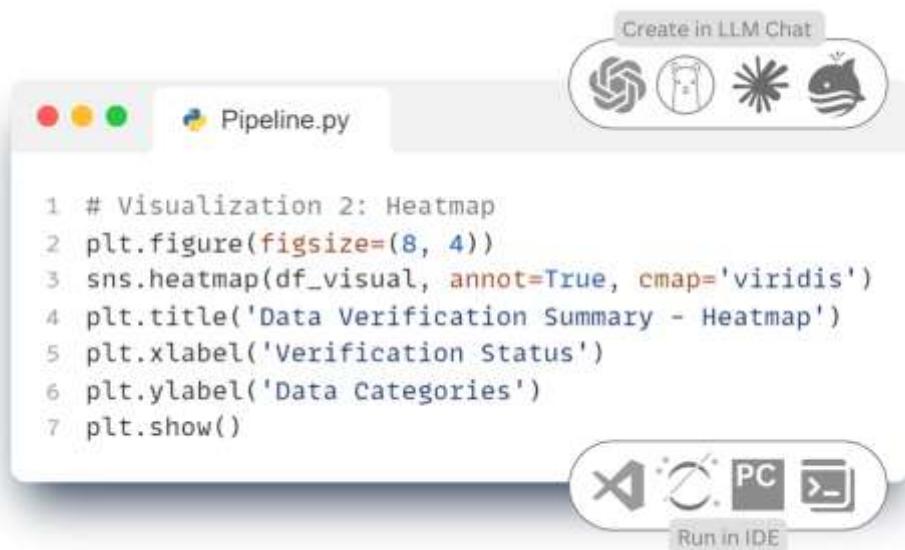
Nav nepieciešamas zināšanas par īpašām vizualizācijas bibliotēkām - mūsdienīgi rīki, tostarp LLM, ļauj automātiski ģenerēt grafiku un veselu lietojumprogrammu kodu, pamatojoties uz uzdevuma aprakstu.

Rīka izvēle ir atkarīga no projekta mērķiem: vai tas ir ziņojums, prezentācija vai tiešsaistes informācijas panelis. Piemēram, atvērtā koda bibliotēka Seaborn ir īpaši laba darbam ar kategoriskiem datiem, palīdzot noteikt modeļus un tendences.

- ☒ Lai redzētu, kā darbojas Seaborn bibliotēka, varat vai nu tieši lūgt LLM izmantot vēlamo bibliotēku, vai arī nosūtīt līdzīgu teksta pieprasījumu LLM uz adresi:

Rādīt karstuma karti ↗

- ☒ LLM atbildi koda un pabeigta grafika veidā, kura attēlošanas kodu tagad var kopēt IDE, bet pašu grafiku var kopēt vai saglabāt, lai to ielīmētu dokumentā:



The screenshot shows a Mac OS X window titled "Pipeline.py". Inside the window, there is a code editor containing the following Python code:

```

1 # Visualization 2: Heatmap
2 plt.figure(figsize=(8, 4))
3 sns.heatmap(df_visual, annot=True, cmap='viridis')
4 plt.title('Data Verification Summary - Heatmap')
5 plt.xlabel('Verification Status')
6 plt.ylabel('Data Categories')
7 plt.show()

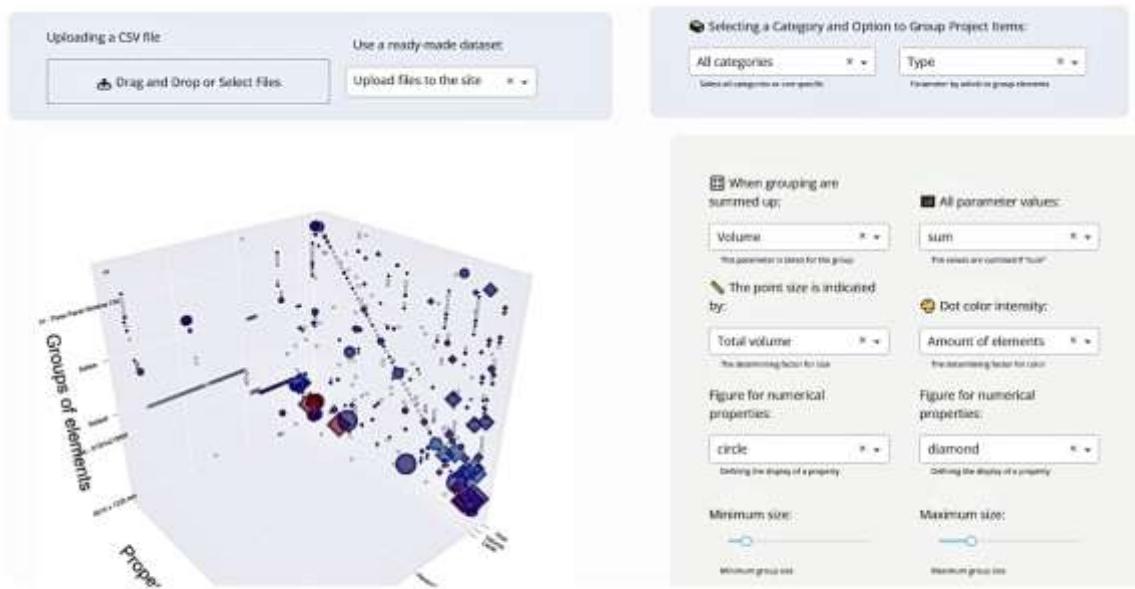
```

At the top right of the window, there is a toolbar with icons for "Create in LLM Chat", "Run in IDE", and other options. Below the toolbar, there are four small circular icons representing different AI models or functions.



**Attēls 7.2-11 Datu validēšanas (attēls 7.2-8) rezultātu vizualizēšana, izmantojot Seaborn bibliotēku.**

Tiem, kas dod priekšroku interaktīvai pieejai, ir pieejami rīki, kas ļauj izveidot dinamiskas diagrammas un paneļus ar iespēju mijiedarboties. Plotly bibliotēka ( 7.1-6. attēls, 7.2-12. attēls) piedāvā iespēju izveidot ļoti interaktīvas diagrammas un paneļus, kurus var iestrādāt tīmekļa lapās un ļaut lietotājam reāllaikā mijiedarboties ar datiem.



**Attēls 7.2-12 Interaktīva 3D elementu atribūtu vizualizācija no CAD- (BIM-) projekta, izmantojot Plotly bibliotēku.**

Specializētās atvērtā pirmkoda bibliotēkas Bokeh, Dash un Streamlit nodrošina ērtu veidu, kā attēlot datus, neprasot padzīlinātas zināšanas tīmekļa izstrādē. Bokeh ir piemēota sarežģītiem interaktīviem grafikiem, Dash izmanto pilnvērtīgu analītisku paneļu veidošanai, bet Streamlit ļauj ātri izveidot tīmekļa lietojumprogrammas datu analīzei.

Izmantojot šādus vizualizācijas rīkus, izstrādātāji un analītiķi var efektīvi izplatīt rezultātus kolēģiem un ieinteresētajām personām, nodrošinot intuitīvu mijiedarbību ar datiem un vienkāršojot lēmumu pieņemšanu.

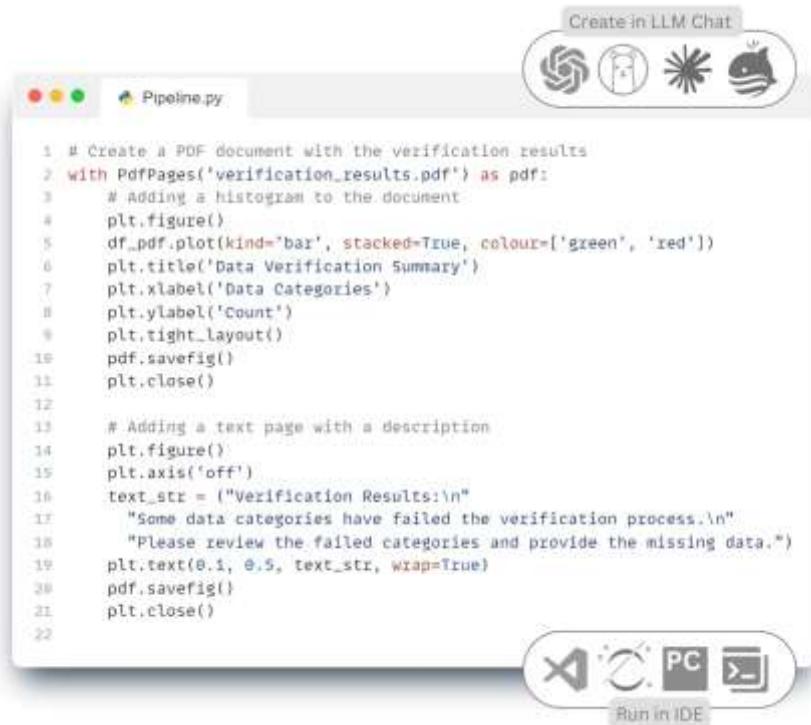
## ETL Load: Automātiska PDF dokumentu izveide.

Datu ielādes posmā varat ne tikai vizualizēt datus, augšupielādēt tos tabulās vai datubāzēs, bet arī automātiski ģenerēt pārskatus, tostarp nepieciešamos grafikus, diagrammas un galvenos analītiskos rādītājus, ko var saņemt vadītājs vai speciālists, kurš gaida revīzijas rezultātus. Automatizētajos pārskatos var iekļaut gan komentārus, gan datu tekstuālu interpretāciju, gan vizualizācijas elementus - tabulas, grafikus.

- 2 Lai izveidotu PDF -ziņojumu ar histogrammu ( 7.2-9. attēls) un analīzes aprakstu, pamatojoties uz iepriekšējās nodalījās veikto validāciju, ir pietiekami formulēt pieprasījumu dialoga turpinājumā ar LLM, piemēram,:

Uzrakstiet kodu, lai izveidotu PDF - failu ar histogrammu un aprakstu par iepriekš (tērzēšanas laikā) veiktās datu validācijas rezultātiem, kā arī uzrakstiet teksta brīdinājumu, ka dažas kategorijas nav validētas un ka trūkstošie ir jāaizpilda. ↵

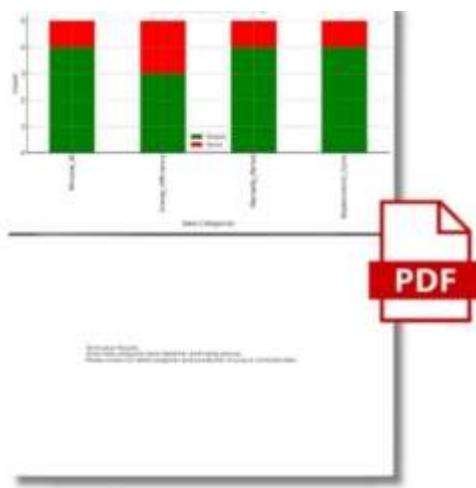
- 2 LLM atbilde kā kods un pabeigts PDF ar rezultātiem:



```

1 # Create a PDF document with the verification results
2 with PdfPages('verification_results.pdf') as pdf:
3     # Adding a histogram to the document
4     plt.figure()
5     df_pdf.plot(kind='bar', stacked=True, colour=['green', 'red'])
6     plt.title('Data Verification Summary')
7     plt.xlabel('Data Categories')
8     plt.ylabel('Count')
9     plt.tight_layout()
10    pdf.savefig()
11    plt.close()
12
13    # Adding a text page with a description
14    plt.figure()
15    plt.axis('off')
16    text_str = ("Verification Results:\n"
17                "Some data categories have failed the verification process.\n"
18                "Please review the failed categories and provide the missing data.")
19    plt.text(0.1, 0.5, text_str, wrap=True)
20    pdf.savefig()
21    plt.close()
22

```



**Attēls 7.2-13 Automatizētais kods izveido PDF -dokumentu, kas satur histogrammu ar testa datiem un tekstu ar testa rezultātiem.**

Automātiski uzrakstīts tikai 20 rindu kods, izmantojot LLM, uzreiz izveido vēlamo PDF (vai DOC) dokumentu ar vizualizāciju atribūtu histogrammas veidā (7.2-13. attēls), kurā parādīts to datu skaits, kas izturējuši un neizturējuši validāciju, kā arī pievienots teksta bloks, kurā apkopoti rezultāti un ieteikumi turpmākai rīcībai.

Automatizēta dokumentu ģenerēšana ir galvenais elements ielādes posmā, jo īpaši projekta vidē, kur ziņošanas ātrums un precizitāte ir ļoti svarīgi.

## ETL Load: automātiska dokumentu ģenerēšana no FPDF

Ziņošanas automatizēšana ETL posmā Load ir svarīgs datu apstrādes posms, jo īpaši tad, ja analīžu rezultāti ir jāsniedz viegli saprotamā un saprotamā formātā. Būvniecības nozarē tas bieži ir svarīgi progresu ziņojumiem, projekta datu statistikai, kvalitātes nodrošināšanas ziņojumiem vai finanšu dokumentācijai.

Viens no ērtākajiem rīkiem šādu uzdevumu veikšanai ir atvērtā pirmkoda bibliotēka, FPDF, kas pieejama gan Python, gan PHP.

Atvērtā pirmkoda **FPDF** bibliotēka nodrošina elastīgu veidu, kā ġenerēt dokumentus, izmantojot kodu, lūujot pievienot galvenes, tekstu, tabulas un attēlus. Koda izmantošana manuālas redīgēšanas vietā samazina kļūdu skaitu un paātrina ziņojumu sagatavošanas procesu PDF formātā.

Viens no galvenajiem soļiem, veidojot PDF -dokumentu, ir virsrakstu un galvenā teksta pievienošana komentāru vai aprakstu veidā. Tomēr, veidojot pārskatu, ir svarīgi ne tikai pievienot tekstu, bet arī to pareizi strukturēt. Virsraksti, atkāpes, atstarpes starp rindām - tas viss ietekmē dokumenta lasāmību. Izmantojot FPDF, varat iestatīt formatēšanas parametrus, kontrolēt elementu izkārtojumu un pielāgot dokumenta stilu.

FPDF principā ir ļoti līdzīgs HTML. Tie, kas jau pārzina HTML, var viegli izveidot jebkuras sarežģītābas PDF dokumentus, izmantojot FPDF, jo koda struktūra ir ļoti līdzīga HTML iezīmēšanai: virsraksti, teksts, attēli un tabulas tiek pievienoti līdzīgā veidā. Tiem, kas nepārzina HTML, nav jāuztraucas - var izmantot LLM, kas uzreiz palīdzēs sastādīt kodu, lai ģenerētu vēlamo dokumenta izkārtojumu.

- ☒ Nākamajā piemērā ir parādīts, kā ģenerēt pārskatu ar galveni un pamattekstu. Izpildot šo kodu jebkurā IDE ar Python atbalstu, tiek izveidots PDF fails, kas satur vēlamo galveni un tekstu:

```
from fpdf importet FPDF      # Importēt FPDF bibliotēku
pdf = FPDF()    # Izveidot PDF -document
pdf.add_page()   # Pievienot lapu

pdf.set_font("Arial", style='B', size=16) # Iestatīt fontu: Arial, treknraksts, izmērs 16
pdf.cell(200, 10, "Projekta ziņojums", ln=True, align='C') # Izveidojiet virsrakstu un
centrējiet to
pdf.set_font("Arial", size=12) # Mainiet fontu uz parasto Arial, izmērs 12
pdf.multi_cell(0, 10, "Šis dokuments satur datus par projekta datnes pārbaudes rezultātiem . . .") #
Pievienojet daudzrindu tekstu
pdf.output(r "C:\reports\report.pdf") # Saglabāt PDF -file
```



**Attēls 7.2-14 Izmantojot dažas Python koda rindas, mēs varam automātiski ģenerēt vajadzīgo PDF teksta dokumentu.**

Sagatavojojot pārskatus, ir svarīgi ņemt vērā, ka dati, no kuriem tiek veidots dokuments, reti kad ir statiski. Virsraksti, teksta bloki ( 7.2-14. attēls) bieži tiek veidoti dinamiski, saņemot vērtības transformācijas posmā ETL procesā.

Izmantojot kodu, varat izveidot dokumentus, kas satur aktuālu informāciju: projekta nosaukumu, pārskata izveides datumu, kā arī informāciju par dalībniekiem vai pašreizējo statusu. Mainīgo lielumu izmantošana kodā ļauj automātiski ievietot šos datus vajadzīgajās ziņojuma vietās, pilnībā novēršot nepieciešamību tos pirms nosūtīšanas rediģēt manuāli.

Papildus vienkāršam tekstam un virsrakstiem īpašu vietu projekta dokumentācijā ieņem tabulas. Gandrīz katrā dokumentā ir strukturēti dati: no objektu aprakstiem līdz pārbaudes rezultātiem. Automātiska tabulu ģenerēšana, pamatojoties uz datiem no transformācijas posma, ļauj ne tikai

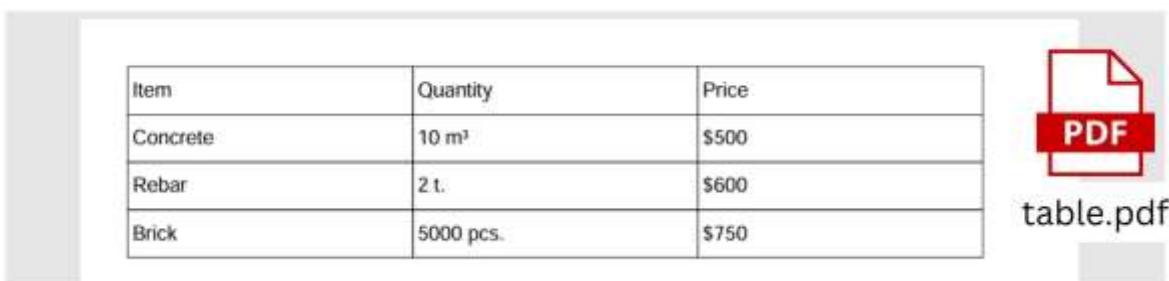
paātrināt dokumentu sagatavošanas procesu, bet arī samazināt kļūdu skaitu, pārnesot informāciju. FPDF ļauj ievietot tabulas PDF failos (kā tekstu vai attēlus), iestatot šūnu robežas, kolonnu izmērus un fontus ( 7.2-15. attēls). Tas ir īpaši ērti, strādājot ar dinamiskiem datiem, kad rindu un kolonnu skaits var mainīties atkarībā no dokumenta uzdevumiem.

- Turpmākajā piemērā parādīts, kā automatizēt tabulu izveidi, piemēram, ar materiālu sarakstu, aplēsēm vai parametru testu rezultātiem:

```
dati = [
    ["Prece", "Daudzums", "Cena"],   # Slezas virsraksti
    ["Betons", "10 m3", "$ 500."], # Pirmās rindas dati
    ["Armatūra", "2 tonnas", "$ 600"], # Otrās rindas dati.
    ["Kieģelis", "5 000 gab.", "$ 750."], # 3. rindas dati.
]

pdf = FPDF () # Izveidot PDF -document
pdf.add_page() # Pievienot lapu
pdf.set_font("Arial", size=12) # Iestatīt fontu

for row in data: # Meklē tabulas rindas
    for item in row: # Pāriet caur rindas šūnām
        pdf.cell(60, 10, item, border=1) # Izveido šūnu ar robežu, platumu 60 un augstumu 10
    pdf.ln() # Pāriet uz nākamo rindu
pdf.output(r "C:\reports\table.pdf") # Saglabāt PDF -file
```



Item	Quantity	Price
Concrete	10 m <sup>3</sup>	\$500
Rebar	2 t.	\$600
Brick	5000 pcs.	\$750



table.pdf

**Attēls 7.2-15 Varat automātiski ģenerēt ne tikai tekstu, bet arī jebkuru tabulas informāciju no transformēšanas soļa PDF.**

Reālajos ziņošanas scenārijos tabulas parasti ir dinamiski ģenerēta informācija, kas iegūta datu pārveidošanas posmā. Dotajā piemērā ( 7.2-15. attēls) tabula ir ievietota PDF -dokumentā statiskā formā: dati šim piemēram tika ievietota datu vārdnīcā (koda pirmā rindiņa), reālos apstākļos šāds datu mainīgais tiek aizpildīts automātiski pēc, piemēram, datu rāmja grupēšanas.

Praksē šādas tabulas bieži tiek veidotas, pamatojoties uz strukturētiem datiem, kas nāk no dažādiem dinamiskiem avotiem: datu bāzēm, Excel failiem, API saskarnēm vai analītisko aprēķinu rezultātiem. Visbiežāk pārveidošanas (ETL) posmā dati tiek apkopoti, grupēti vai filtrēti un tikai pēc tam pārveidoti kopsummās grafiku vai divdimensiju tabulu veidā, kas tiek parādītas pārskatos. Tas nozīmē, ka tabulu saturs var mainīties atkarībā no izvēlētajiem parametriem, analīzes perioda, projekta filtriem vai lietotāja iestatījumiem.

Dinamisko datu rāmju un datu kopu izmantošana transformēšanas posmā padara ziņošanas procesu ielādes posmā pēc iespējas elastīgāku, mērogojamu un viegli atkārtojamu bez nepieciešamības veikt manuālu iejaukšanos.

Papildus tabulām un tekstam FPDF atbalsta arī tabulāro datu grafiku pievienošanu, kas ļauj ziņojumā ievietot attēlus, kas ģenerēti ar Matplotlib vai citām iepriekš aplūkotajām vizualizācijas bibliotēkām. Jebkurus grafikus, diagrammas un diagrammas var pievienot dokumentam, izmantojot kodu.

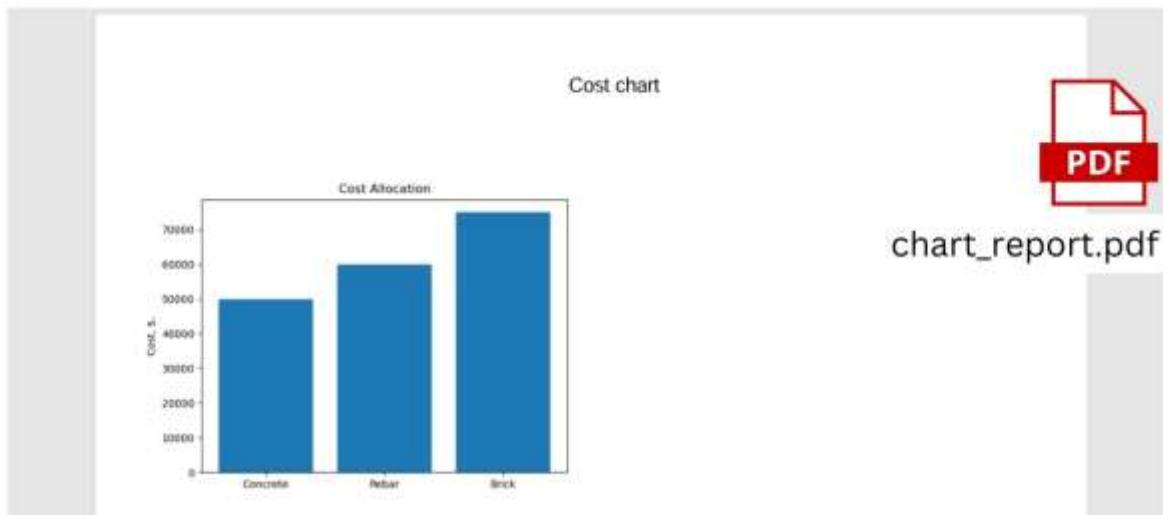
- 2 Izmantojot Python bibliotēku FPDF, pievienosim PDF dokumentam grafiku, kas iepriekš ģenerēts ar Matplotlib:

```
import matplotlib.pyplot as plt # Import Matplotlib, lai izveidotu grafikus

fig, ax = plt.subplots() # Izveidot grafika figūru un asis
categories = ["Concrete", "Rebar", "Brick"] # Kategoriju nosaukumi
values = [50000, 60000, 75000] # Kategorijas vērtības
ax.bar(categories, values) # Izveidot joslu diagrammu
plt.ylabel("Value,$.") # Parakstiet Y asi
plt.title("Izmaksu sadalījums") # Pievienot virsrakstu
plt.savefig(r "C:\reports\chart\chart\chart\chart.png") # Saglabāt diagrammu kā attēlu

pdf = FPDF () # Izveidot PDF -dokuments
pdf.add_page() # Pievienot lapu
pdf.set_font("Arial", size=12) # Iestatīt fontu
pdf.cell(200, 10, "Cost Chart", ln=True, align="C") # Pievieno galveni

pdf.image(r "C:\reports\chart\chart\chart\chart.png", x=10, y=30, w=100) # Ievietojiet attēlu PDF failā (x, y - koordinātas, w - platums)
pdf.output(r "C:\reports\chart_report.pdf") # Saglabājiet PDF failu
```



**Attēls 7.2-16 Izmantojot tikai duci kodu rindu, varat ģenerēt grafiku, saglabāt to un pēc tam ielīmēt PDF dokumentā.**

Ar FPDF dokumentu sagatavošanas un loģikas process kļūst pārskatāms, ātrs un ērts. Kodā iebūvētās veidnes ļauj ģenerēt dokumentus ar aktuāliem datiem, novēršot nepieciešamību tos aizpildīt manuāli.

Izmantojot ETL automatizāciju - laikietilpīgas manuālas atskaišu sagatavošanas vietā speciālisti var koncentrēties uz datu analīzi un lēmumu pieņemšanu, nevis izvēlēties pareizo rīku darbam ar konkrētu datu bunkuru ar skaidru lietotāja saskarni.

Tādējādi FPDF bibliotēka nodrošina elastīgu rīku jebkuras sarežģītības dokumentu automatizētai izveidei - no ūsiem tehniskiem ziņojumiem līdz sarežģītiem analītiskiem kopsavilkumiem ar tabulām un diagrammām, kas ļauj ne tikai paātrināt dokumentu plūsmu, bet arī ievērojami samazināt kļūdu iespējamību, kas saistīta ar manuālu datu ievadīšanu un formatēšanu.

### **ETL Load: Ziņošana un ielāde uz citām sistēmām.**

Iekraušanas posmā rezultāti tika sagatavoti tabulu, grafiku un galīgo PDF ziņojumu veidā, kas sagatavoti saskaņā ar noteiktajām prasībām. Turpmāk šos datus ir iespējams eksportēt mašīnlasāmos formātos (piemēram, CSV), kas nepieciešami integrācijai ar ārējām sistēmām, piemēram, ERP, CAFM, CPM, BI platformām un citiem uzņēmuma vai nozares risinājumiem. Papildus CSV formātam datus var augšupielādēt XLSX, JSON, XML vai tieši datu bāzēs, kas atbalsta automātisku informācijas apmaiņu.

- ❷ Lai ģenerētu atbilstošu kodu Load soļa automatizēšanai, vienkārši pieprasiet LLM - interfeisu, piemēram: ChatGPT, LlaMa, Mistral DeepSeek, Grok, Claude vai QWEN:

Uzrakstiet kodu, lai ģenerētu datu validācijas rezultātu pārskatu datu rāmī, kurā slejas ar prefiksū "verified\_" tiek saskaitītas, pārdēvētas par "Passed" un "Failed", trūkstošās vērtības tiek aizstātas ar 0, un pēc tam tikai tās rindas, kas iztur visas validācijas, tiek eksportētas uz CSV - failu.

 LLM atbilde:



```

1 # Create a report on the results of data validation
2 report = df.filter(like='verified_').apply(lambda x: x.value_counts());T
3 report.columns = ['Passed', 'Failed']
4 report.fillna(0, inplace=True)
5
6 # Export the cleaned and verified data to a CSV file
7 verified_df = df[df.filter(like='verified_').all(axis=1)]
8 verified_df.to_csv('verified_data.csv', index=False)

```

**Attēls 7.2-17 Pārveidošanas posmā iegūtos validētos datus no galīgās datu kopas eksportē CSV failā, lai integrētu ar citām sistēmām.**

Iepriekš minētais kods ( 7.2-17. attēls) īsteno ETL -procesa pēdējo posmu - Load (ielāde), kura laikā pārbaudītie dati tiek saglabāti CSV formātā, kas ir saderīgs ar lielāko daļu ārējo sistēmu un datubāzu. Tādējādi mēs esam pabeiguši pilnu ETL procesa ciklu, ieskaitot datu ieguvi, transformāciju, vizualizāciju, dokumentēšanu un eksportēšanu uz mums vajadzīgajām sistēmām un formātiem, kas nodrošina atkārtojamību, pārredzamību un darba ar informāciju automatizāciju.

ETL cauruļvadu var izmantot gan atsevišķu projektu apstrādei, gan liela mēroga lietojumprogrammām, analizējot simtiem vai tūkstošiem ienākošo datu dokumentu, attēlu, skenējumu, CAD projektu, punktu mākoņu, PDF failu vai citu avotu veidā, kas nāk no izplatītām sistēmām. Iespēja pilnībā automatizēt procesu padara ETL ne tikai par tehnisku apstrādes rīku, bet arī par digitālās būvniecības informācijas infrastruktūras pamatu.

## ETL ar LLM: datu vizualizēšana no PDF -dokumentiem

Ir pienācis laiks pāriet pie pilnvērtīga ETL procesa izveides, kas aptver visus galvenos datu apstrādes posmus vienā scenārijā - ieguvi, pārveidošanu un ielādi. Izveidosim automatizētu ETL cauruļvadu, kas ļauj apstrādāt PDF dokumentus bez manuāla darba - iegūt datus no dokumentiem, vizualizēt, analizēt un pārsūtīt uz citām sistēmām.

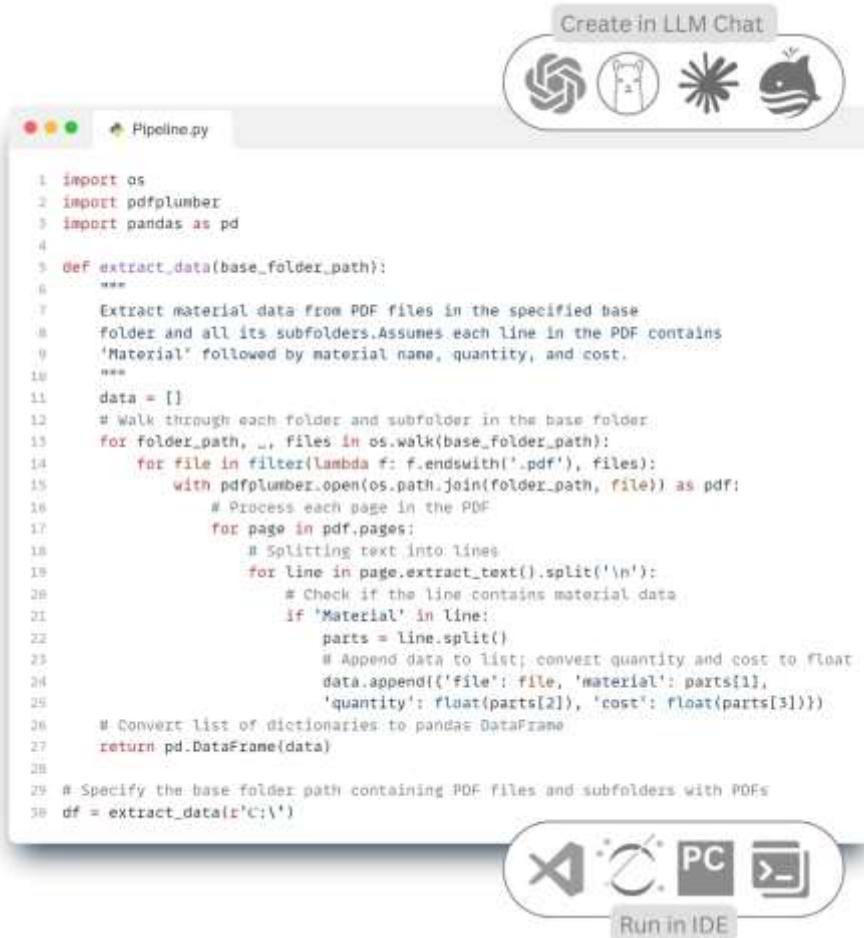
Mūsu piemērā ETL process tiks aprakstīts, izmantojot uzvednes, kurām valodas modelim (LLM) būs

jāizskaidro visi ETL procesi ar iegūstamā gala rezultāta aprakstu. Šajā gadījumā uzdevums ir atrast visus PDF failus norādītajā mapē un tās apakšmapēs, iegūt no tiem attiecīgo informāciju, piemēram, materiālu nosaukumus, daudzumus un izmaksas, un iesniegt rezultātu kā strukturētu tabulu (DataFrame) turpmākai analīzei

- 💡 Pirmais teksta vaicājums LLM, lai automātiski izvilktu datus no vairākiem PDF dokumentiem un izveidotu datu rāmi datu izvilkšanas solim:

Uzrakstiet kodu, lai iegūtu materiālu informāciju no PDF failiem noteiktā mapē un tās apakšmapēs. PDF datnēs ir ietverti dati par materiāla nosaukumu, daudzumu un izmaksām. Rezultāts jāglabā datu rāmī.

- 💡 LLM atbilde :



```

1 import os
2 import pdfplumber
3 import pandas as pd
4
5 def extract_data(base_folder_path):
6     """
7         Extract material data from PDF files in the specified base
8         folder and all its subfolders. Assumes each line in the PDF contains
9         'Material' followed by material name, quantity, and cost.
10    """
11    data = []
12    # Walk through each folder and subfolder in the base folder
13    for folder_path, _, files in os.walk(base_folder_path):
14        for file in filter(lambda f: f.endswith('.pdf'), files):
15            with pdfplumber.open(os.path.join(folder_path, file)) as pdf:
16                # Process each page in the PDF
17                for page in pdf.pages:
18                    # Splitting text into lines
19                    for line in page.extract_text().split('\n'):
20                        # Check if the line contains material data
21                        if 'Material' in line:
22                            parts = line.split()
23                            # Append data to list; convert quantity and cost to float
24                            data.append({'file': file, 'material': parts[1],
25                                         'quantity': float(parts[2]), 'cost': float(parts[3])})
26    # Convert list of dictionaries to pandas DataFrame
27    return pd.DataFrame(data)
28
29 # Specify the base folder path containing PDF files and subfolders with PDFs
30 df = extract_data(r'C:\')

```

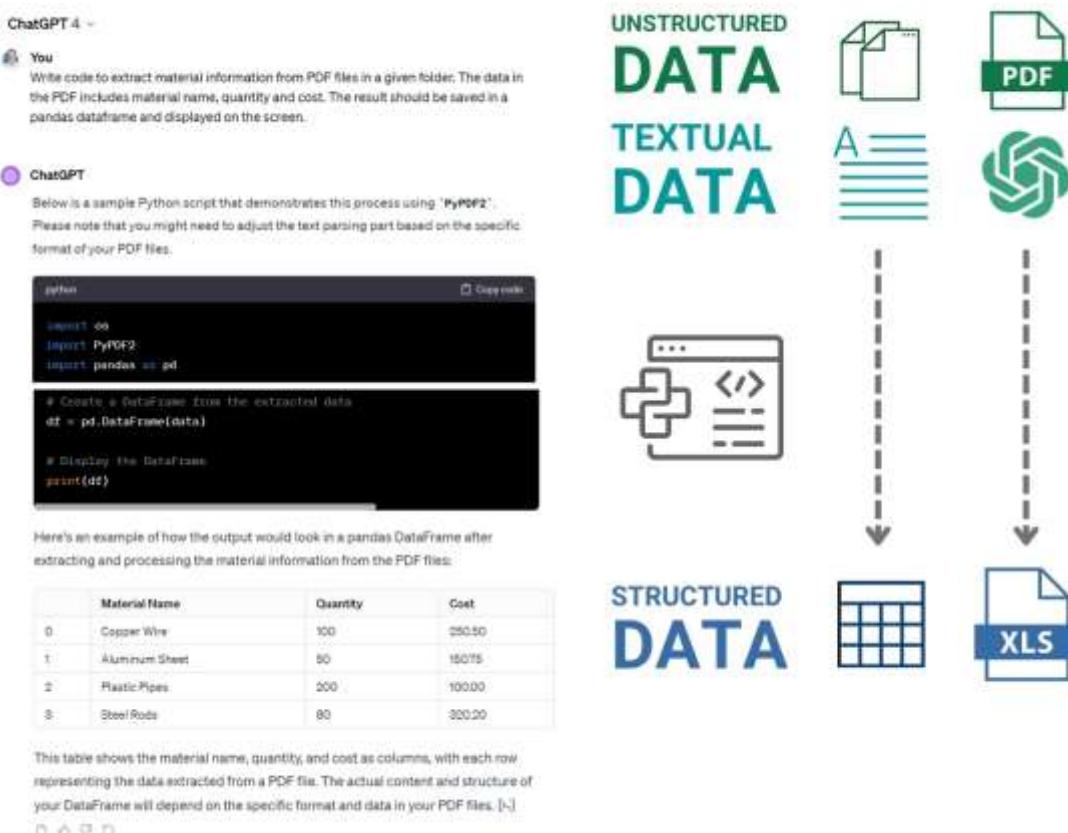
The screenshot shows a code editor window titled "Pipeline.py" containing the provided Python script. Above the editor is a toolbar with icons for "Create in LLM Chat" and four AI models: a spiral icon, a cat icon, a starburst icon, and a whale icon. Below the editor is another toolbar with icons for "Run in IDE" (Visual Studio Code), "PC", and a terminal or file icon.

Attēls 7.2-18 LLM izveido Python kodu, lai iegūtu datus no PDF failiem noteiktā mapē un visās tās apakšmapēs.

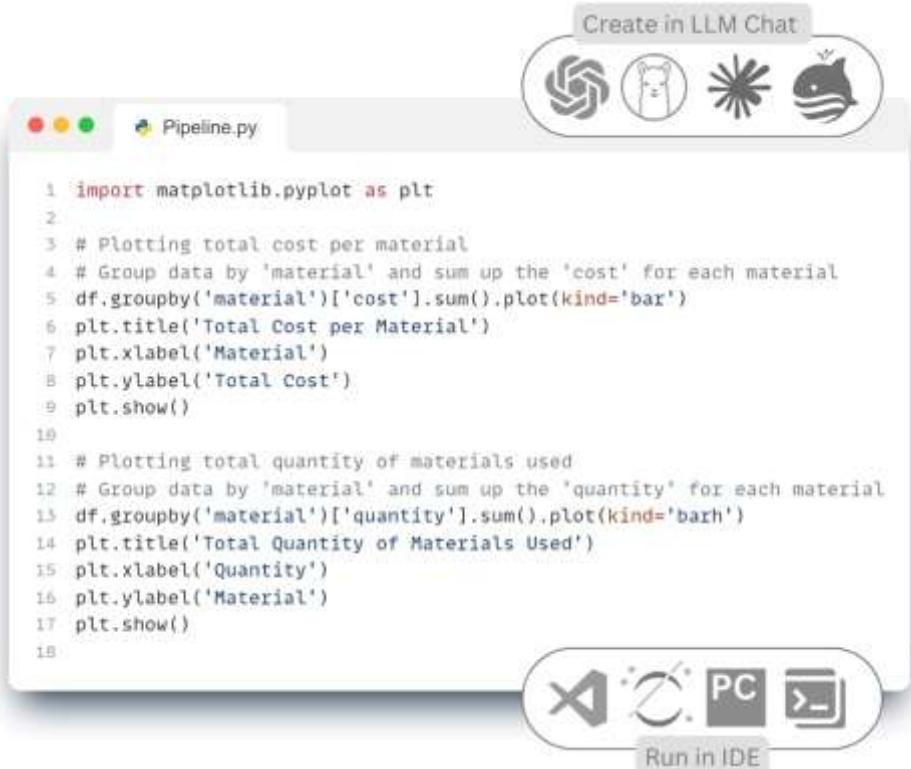
LLM atbilde ( 7.2-18. attēls) ir gatavs Python -skripts, kas automātiski šķērso visas mapes, atver

atrastos PDF -failus, izraksta no tiem teksta informāciju un pārvērš to tabulā. Iegūto tērzēšanas kodu var palaist kādā no populārākajām IDE PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populāriem tiešsaistes rīkiem Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

Alternatīvi, lai vienkāršotu vākšanas procesu, tā vietā, lai kopētu kodu no LLM un izmantotu kodu IDE, mēs varam arī augšupielādēt duci PDF -failu tieši LLM tērzēšanā ( 7.2-19. attēls) un kā rezultātu saņemt tabulu, nerēdzot kodu un to neskrienot. Šī koda izpildes rezultāts būs tabula ar mūsu izvēlētajiem atribūtiem.



Vizualizējet katra materiāla kopējās izmaksas un daudzumu no datu rāmja ( 7.2-18. attēls) ↴



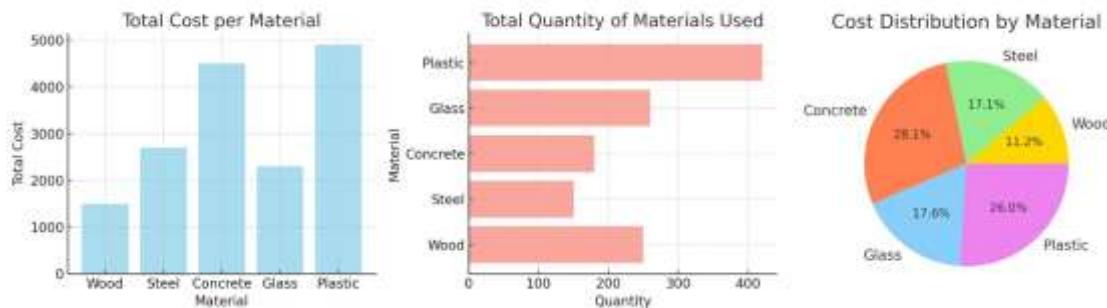
```

1 import matplotlib.pyplot as plt
2
3 # Plotting total cost per material
4 # Group data by 'material' and sum up the 'cost' for each material
5 df.groupby('material')['cost'].sum().plot(kind='bar')
6 plt.title('Total Cost per Material')
7 plt.xlabel('Material')
8 plt.ylabel('Total Cost')
9 plt.show()
10
11 # Plotting total quantity of materials used
12 # Group data by 'material' and sum up the 'quantity' for each material
13 df.groupby('material')['quantity'].sum().plot(kind='barh')
14 plt.title('Total Quantity of Materials Used')
15 plt.xlabel('Quantity')
16 plt.ylabel('Material')
17 plt.show()
18

```

Attēls 7.2-20 LLM atbildē -modeli kā Python kods, lai vizualizētu datus no datu rāmja, izmantojot matplotlib bibliotēku.

LLM automātiski ģenerē un izpilda Python kodu ( 7.2-20. attēls), izmantojot matplotlib bibliotēku. Pēc šī koda izpildes mēs iegūstam izmaksu un materiālu izmantošanas grafikus būvniecības projektos tieši čatā ( 7.2-21. attēls), kas ievērojami atvieglo analītisko darbu.

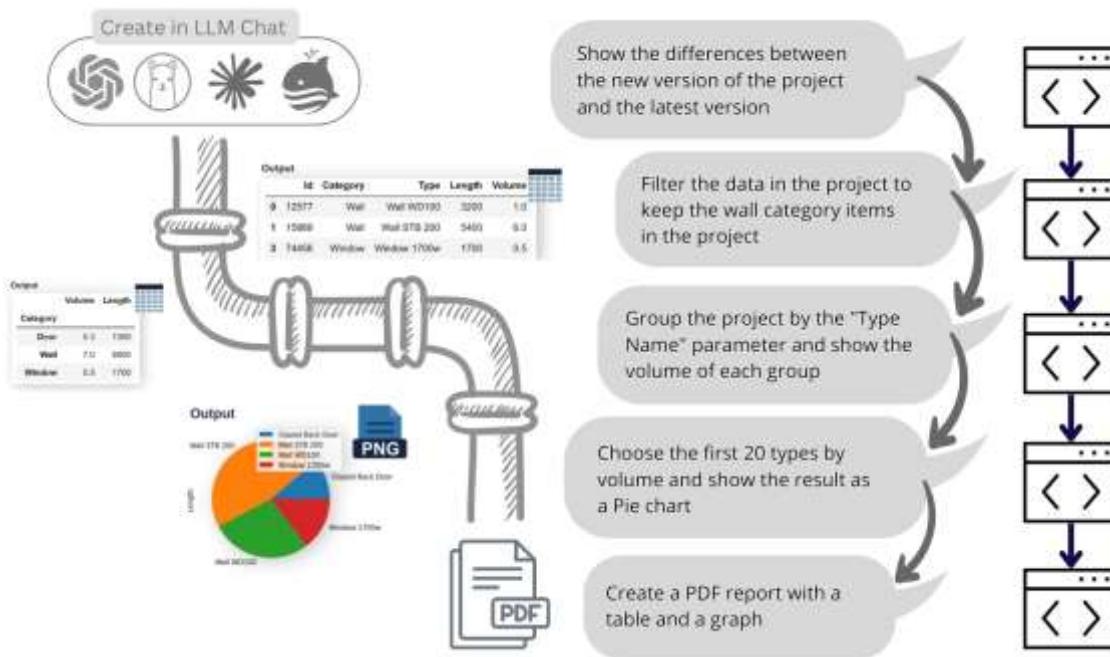


Attēls 7.2-21 LLM atbildes vizualizācija grafiku veidā, pamatojoties uz datiem, kas apkopoti datu rāmī.

Atbalsts ideju izstrādē par ETL koda rakstīšanu, koda analīzi un izpildi, kā arī rezultātu vizualizēšanu ir

pieejams, izmantojot vienkāršus teksta pieprasījumus LLM, un nav nepieciešams apgūt programmēšanas pamatus. Mākslīgā intelekta rīku, piemēram, LLM, parādīšanās noteikti maina pieeju programmēšanai un datu apstrādes automatizēšanai ( 7.2-22. attēls).

Saskaņā ar PwC ziņojumu "Kāda ir mākslīgā intelekta patiesā vērtība jūsu uzņēmumam un kā jūs varat to izmantot?" (2017) [139], procesu automatizācija un produktivitātes uzlabojumi būs galvenie ekonomikas izaugsmes virzītājspēki. Un paredzams, ka produktivitātes uzlabojumi laikposmā no 2017. līdz 2030. gadam veidos vairāk nekā 55% no visa mākslīgā intelekta izraisītā IKP pieauguma".



Attēls 7.2-22 AI LLM palīdz ģenerēt koda projektu, kas tiek izmantots turpmākajos projektos bez nepieciešamības pēc LLM.

Izmantojot tādus rīkus kā ChatGPT, LLaMa, Mistral, Claude, DeepSeek, QWEN, Grok, kā arī atvērtos datus un atvērtā pirmkoda programmatūru, mēs varam automatizēt procesus, kurus iepriekš varēja veikt tikai ar specializētām, dārgām un grūti kopjamām modulārām patentētām sistēmām.

Būvniecības kontekstā tas nozīmē, ka uzņēmumi, kas pirmie ieviešs automatizētus cauruļvadu -datu procesus, gūs ievērojamus ieguvumus, sākot no uzlabotas projektu pārvaldības efektivitātes līdz finansiālo zaudējumu samazināšanai un fragmentētu lietojumprogrammu un atsevišķu datu noliktavu likvidēšanai.

Aprakstītā biznesa uzdevumu izpildes loģika ETL procesā ir būtiska analītikas un datu apstrādes procesu automatizācijas daļa, kas ir plašāka jēdziena - cauruļvadu (Pipelines) - specifiska variācija.

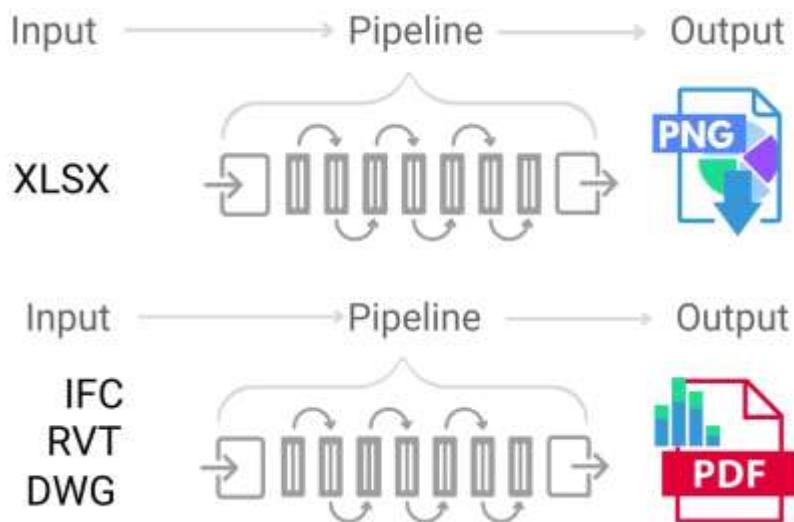
## NODAĻA 7.3.

### AUTOMĀTISKAIS ETL KONVEIJERS (CAURUĻVADS)

#### Cauruļvads: Automātiskais ETL konveijers dati

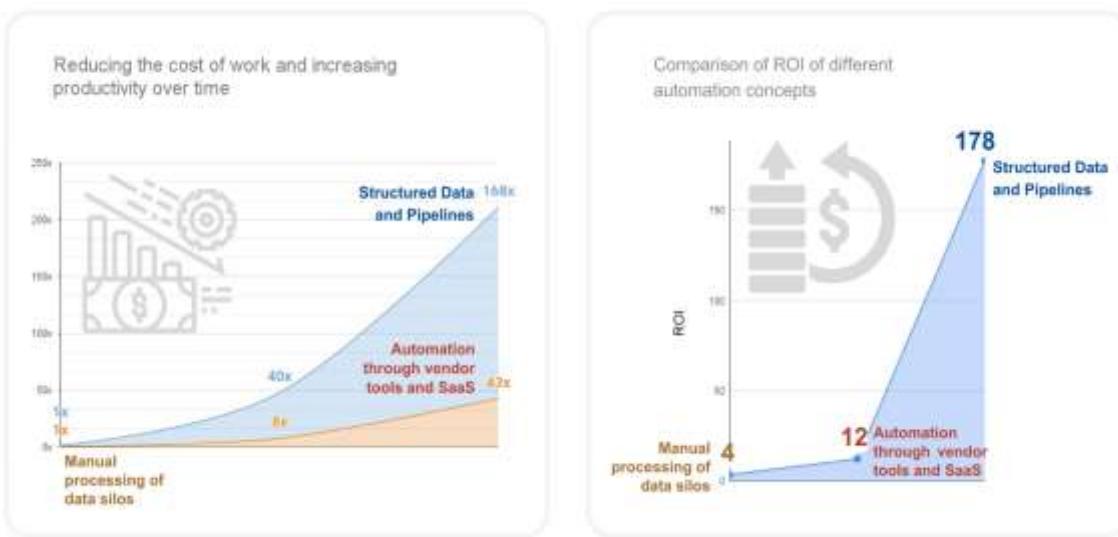
ETL process tradicionāli tiek izmantots datu apstrādei analītiskajās sistēmās, aptverot gan strukturētus, gan nestukturētus avotus. Tomēr mūsdienā digitālajā vidē arvien biežāk tiek lietots plašāks termins - cauruļvads (konveijers), kas apzīmē jebkuru secīgu apstrādes kēdi, kurā viena posma rezultāti kļūst par nākamā posma ievaddatiem.

Šī pieeja attiecas ne tikai uz datiem, bet arī uz citiem automatizācijas veidiem: uzdevumu apstrādi, pārskatu veidošanu, integrāciju ar programmatūru un digitālo darba plūsmu (7.3-1. attēls).



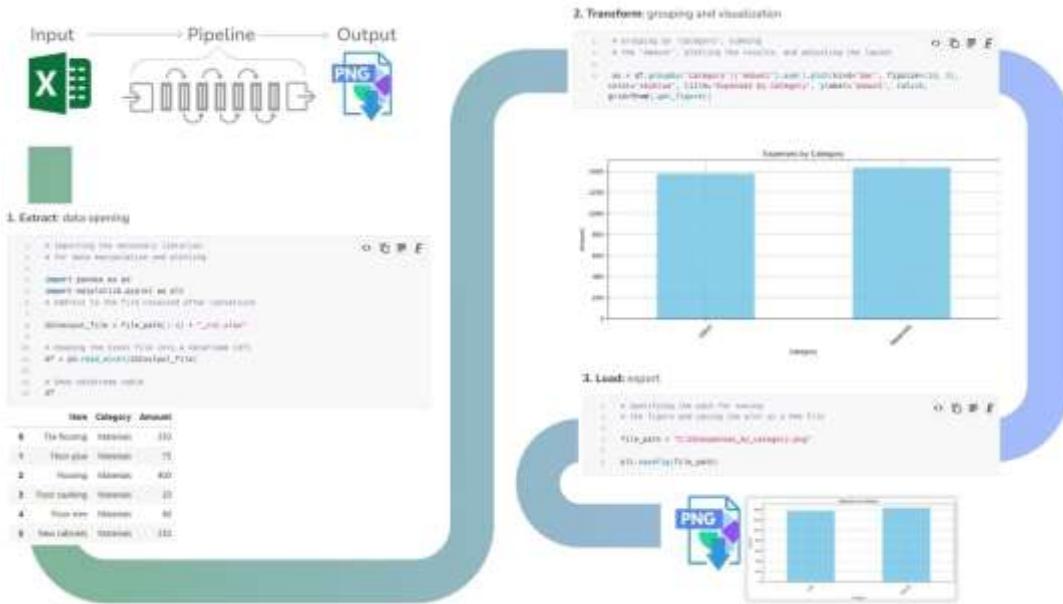
Attēls 7.3-1 Cauruļvads ir apstrādes secība, kurā viena posma izvads kļūst par nākamā posma ievadu.

Cauruļvadu izmantošana ir viens no galvenajiem automatizācijas elementiem, jo īpaši, strādājot ar lielu daudzumu heterogēnu datu. Cauruļvadu arhitektūra ļauj sarežģītus apstrādes soļus organizēt modulārā, konsekventā un pārvaldāmā formātā, kas palielina lasāmību, vienkāršo koda uzturēšanu un ļauj veikt pakāpenisku atkļūdošanu un mērogojamu testēšanu.



**Attēls 7.3-2.** ROI Cauruļvadu datu validācijas process samazina izpildes laiku par desmitiem un simtiem reižu salīdzinājumā ar apstrādi, izmantojot klasiskos rīkus [74].

Atšķirībā no manuāla darba patentētās sistēmās (ERP, PMIS, CAD u.c.), cauruļvads ļauj ievērojamī (7.3-2. attēls) palielināt uzdevumu izpildes ātrumu, izvairīties no atkārtota darba un automatizēt procesu uzsākšanu īstajā laikā (7.3-3. attēls).

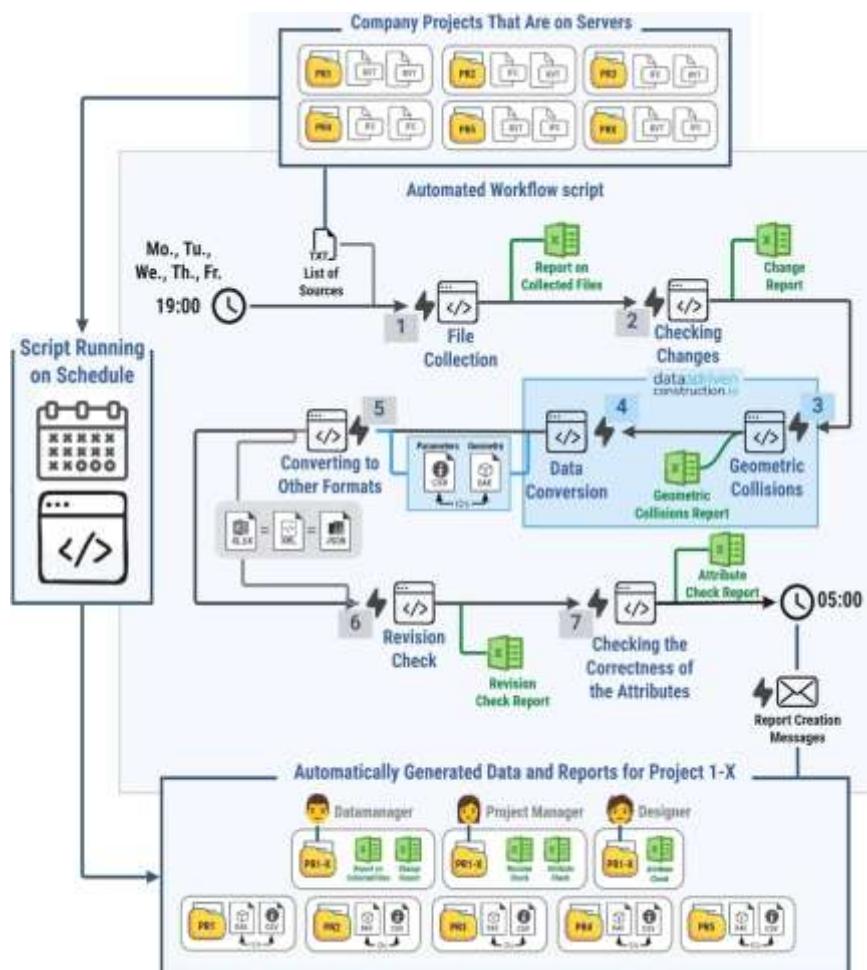


**Attēls 7.3-3** ETL piemērs Cauruļvads, lai automātiski iegūtu grafiku no tabulas datiem XLSX failā bez Excel atvēršanas.

Lai apstrādātu datu plūsmas un izveidotu automatizētu cauruļvadu, līdzīgi kā ETL procesā, ir nepieciešams iepriekš noteikt datu avotus, kā arī to vākšanas laika grafiku - vai nu konkrētam uzņēmējdarbības procesam, vai arī visam uzņēmumam.

Būvniecības projektos dati nāk no daudziem neviendabīgiem avotiem ar dažādiem atjaunināšanas intervāliem. Lai izveidotu uzticamu datu vitrīnu, ir ļoti svarīgi reģistrēt, kad informācija tiek iegūta un atjaunināta. Tas ļauj savlaicīgi pieņemt lēmumus un uzlabo projekta pārvaldības efektivitāti.

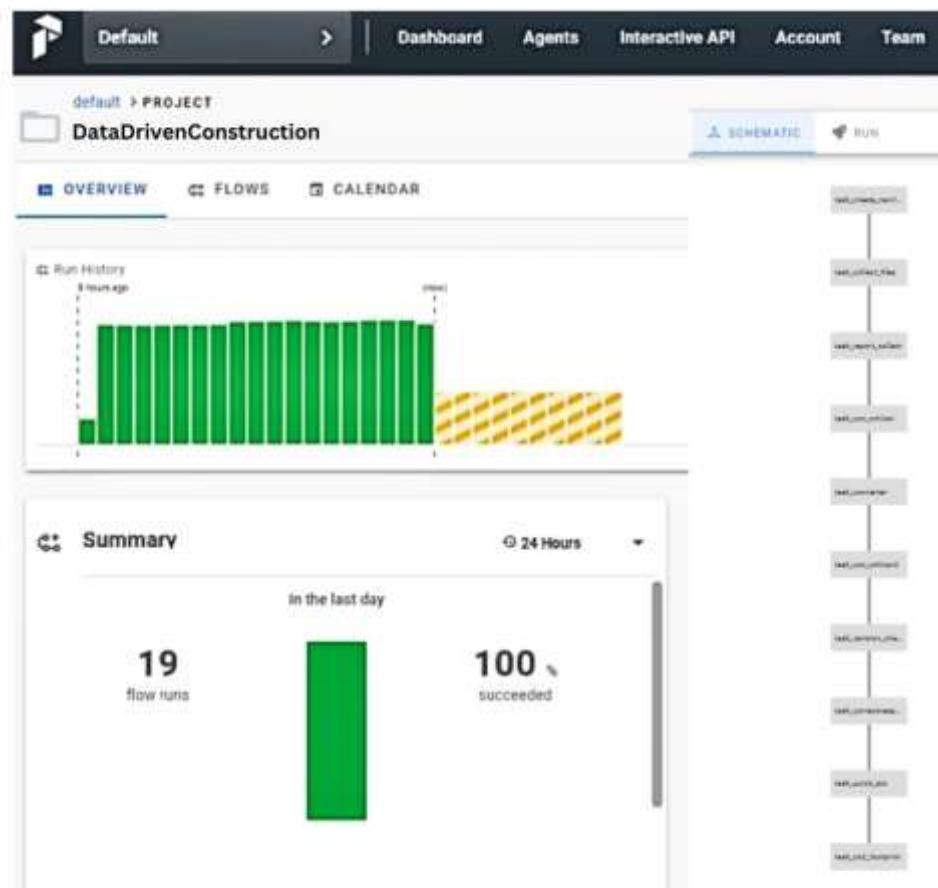
Viena no iespējām ir montāžas procesu sākt noteiktā laikā, piemēram, plkst. 19.00, darba dienas beigās. Šajā brīdī tiek aktivizēts pirmais skripts, kas atbild par datu apkopošanu no dažādām sistēmām un krātuvēm ( 7.3-4. attēls, 1. solis). Pēc tam seko automātiska datu apstrāde un pārveidošana analītikai piemērotā strukturētā formātā ( 7.3-4. attēls, 2-4. solis). Pēdējā posmā, izmantojot sagatavotos datus, tiek automātiski generēti pārskati, paneļi un citi iepriekšējās nodaļās aprakstītie produkti ( 7.3-4. attēls, 6.-7. solis). Rezultātā līdz plkst. 05:00 no rīta vadītājiem jau ir pieejami aktuāli projekta statusa pārskati vajadzīgajā formātā ( 7.3-5. attēls).



Attēls 7.3-4 Dati cauruļvadā, kas automātiski tiek apkopoti vakarā, tiek apstrādāti pa nakti, lai vadītājiem līdz rītam būtu pieejami aktuāli pārskati un svaigi ziņojumi.

Savlaicīga datu vākšana, KPI definēšana, transformācijas procesu automatizācija un vizualizācija, izmantojot informācijas paneļus, ir galvenie elementi veiksmīgai uz datiem balstītu lēmumu pieņemšanai.

Šādus automatizētus procesus ( 7.3-4. attēls) var izpildīt pilnīgi autonomi: tie darbojas pēc grafika, apstrādā datus bez uzraudzības un var tikt izvietoti vai nu mākonī, vai uz uzņēmuma servera ( 7.3-5. attēls). Tas ļauj šādus ETL cauruļvadus integrēt esošajā IT infrastruktūrā, saglabājot kontroli pār datiem un nodrošinot elastību mērogošanas jomā.



Attēls 7.3-5 Automātiskie ETL -konveijera procesi (attēls 7.3-4) Prefect platformā, kurā katru darba dienu pēc plkst. 19:00 pārmaiņus tiek palaisti 10 python skripti.

Darba plūsmu automatizēšana ne tikai palielina komandas produktivitāti, atbrīvojot laiku jēgpilnākiem un mazāk rutīnas uzdevumiem, bet arī kalpo kā svarīgs solis celā uz mākslīgā intelekta (AI) tehnoloģiju iekļaušanu biznesa procesos, ko sīkāk aplūkosim nodaļā par prognozēšanas analīzi un mašīnmācīšanos.

## Cauruļvads -ETL datu validācijas process ar LLM

Iepriekšējās nodalījās par datu prasību izveidi un ETL automatizēšanu, mēs soli pa solim sadalījām datu sagatavošanas, transformēšanas, validēšanas un vizualizācijas procesu. Šīs darbības tika īstenotas kā atsevišķi koda bloki ( 7.2-18. attēls - 7.2-20. attēls), no kuriem katrs veic konkrētu uzdevumu.

Tagad mums ir nākamais mērķis - apvienot šos elementus vienotā, saskaņotā un automatizētā datu apstrādes cauruļvadā - cauruļvadā, ETL -Pipeline -, kurā visi posmi (ielāde, validācija, vizualizācija, eksports) tiek izpildīti secīgi vienā automātiski izpildāmā skripta ietvaros.

Turpmākajā piemērā tiks īstenots pilns datu apstrādes cikls: no avota CSV faila ielādes →līdz struktūras un vērtību pārbaudei, izmantojot regulārās izteiksmes →rezultātu aprēķināšanai→ vizuāla pārskata ģenerēšana PDF formātā.

💡 Lai iegūtu attiecīgo kodu, varat izmantot šādu teksta vaicājumu LLM:

Lūdzu, uzrakstiet koda paraugu, kas ielādē datus no CSV, validē DataFrame, izmantojot regulārās izteiksmes, pārbauda identifikatorus formātā "W-NEW" vai "W-OLD", energoefektivitāti ar burtiem no "A" līdz "G", garantijas periodu un nomaiņas ciklu ar skaitliskajām vērtībām gados un beigās izveido ziņojumu ar pozitīvo un negatīvo vērtību skaitu, ġenerē PDF ar rezultātu histogrammu un pievieno teksta aprakstu. ↴

2 LLM atbilde :

```

1 # Importing necessary libraries
2 import pandas as pd
3 import re
4 import matplotlib.pyplot as plt
5 from matplotlib.backends.backend_pdf import PdfPages
6
7
8 # Load data
9 df = pd.read_csv('path_to_csv_file.csv')
10
11 # Define patterns and verify data
12 patterns = {
13     'Window_ID': r'^W-NEW|W-OLD[0-9]+|D-1[0-9]{2}$',
14     'Energy_Efficiency': r'^[A-G]\+\$',
15     'Warranty_Period': r'^\d+\syears\$',
16     'Replacement_Cycle': r'^\d+\syears\$'
17 }
18
19 for column, pattern in patterns.items():
20     df['verified_' + column] = df[column].apply(lambda x: bool(re.match(pattern, str(x))))
21
22 # Create report
23 report = df.filter(like='verified_').apply(lambda x: x.value_counts()).T
24 report.columns = ['Passed', 'Failed']
25
26 # Generate PDF
27 pdf_path = '/mnt/data/verification_results_full_pipeline.pdf'
28 with PdfPages(pdf_path) as pdf:
29     # Plot and save bar chart
30     report.plot(kind='bar', stacked=True, color=['green', 'red'])
31     plt.title('Data Verification Summary')
32     plt.xlabel('Data Categories')
33     plt.ylabel('Count')
34     plt.tight_layout()
35     pdf.savefig()
36     plt.close()
37
38     # Add text page
39     plt.figure()
40     plt.axis('off')
41     text_str = "Verification Results Summary"
42     plt.text(0.1, 0.5, text_str, wrap=True)
43     pdf.savefig()
44     plt.close()
45
46 pdf_path
47

```

Attēls: 7.3-6 Cauruļvads (ETL) automatizē pilnu datu apstrādes ciklu: no lejupielādes un validācijas līdz strukturēta pārskata izveidei PDF formātā .

Automatizētais kods (7.3-6. attēls) LLM tērzēšanas telpā vai DIE pēc koda kopēšanas pārbaudīs datus no CSV -faila, izmantojot norādītās regulārās izteiksmes, izveidos ziņojumu par pieņemto un noraidīto ierakstu skaitu un pēc tam saglabās pārbaudes rezultātus kā PDF -lietu.

Šāda ETL -konveijera struktūra, kurā katrs posms - no datu ielādes līdz pārskatu ģenerēšanai - tiek īstenots kā atsevišķs modulis, nodrošina pārredzamību, mērogojamību un atkārtojamību. Validācijas loģikas attēlošana kā viegli lasāms Python kods padara procesu pārredzamu un saprotamu ne tikai izstrādātājiem, bet arī datu pārvaldības, kvalitātes un analītikas speciālistiem.

Pipeline pieeja datu apstrādes automatizēšanai ļauj standartizēt procesus, palielināt to atkārtojamību un vienkāršot pielāgošanu jauniem projektiem. Tādējādi tiek radīta vienota datu analīzes metodoloģija neatkarīgi no datu avota vai uzdevuma veida - neatkarīgi no tā, vai tā ir atbilstības pārbaude, pārskatu sniegšana vai datu pārsūtīšana uz ārējām sistēmām.

Šāda automatizācija samazina cilvēcisko kļudu skaitu, mazina atkarību no patentētiem risinājumiem un palielina rezultātu precīzitāti un uzticamību, padarot tos piemērotus gan operatīvajai analīzei projektu līmenī, gan stratēģiskajai analīzei uzņēmuma līmenī.

## Caurulvads -ETL: projekta elementu datu un informācijas pārbaude CAD (BIM).

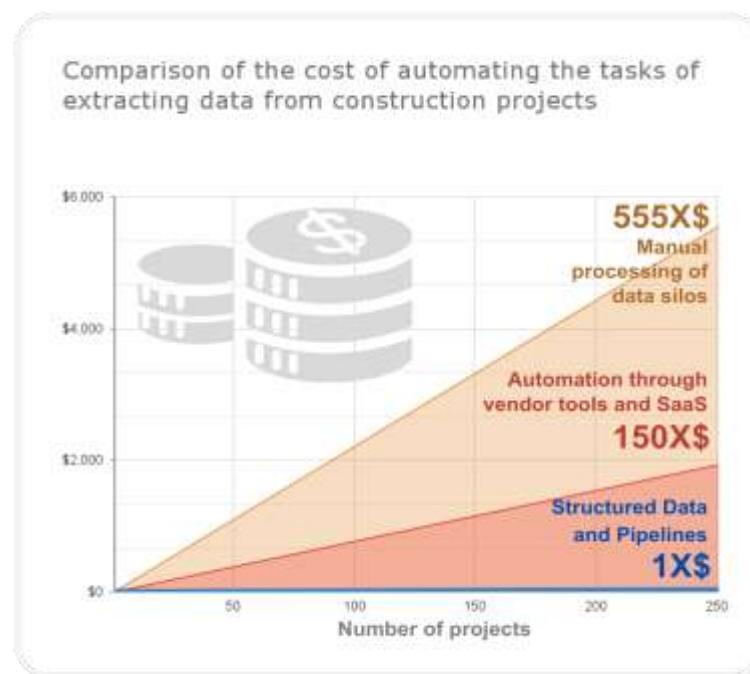
Dati no CAD sistēmām un datubāzēm (BIM) ir vieni no sarežģītākajiem un dinamiskāk atjauninātajiem datu avotiem būvniecības uzņēmumu darbībā. Šīs lietojumprogrammas ne tikai apraksta projektu, izmantojot ģeometriju, bet arī papildina to ar vairākiem teksta informācijas slāniem: apjomiem, materiālu īpašībām, telpu iedalījumu, energoefektivitātes līmeņiem, pielaidēm, dzīves cikliem un citiem raksturlielumiem.

CAD -modeļos vienībām piešķirtie atribūti tiek veidoti projektēšanas posmā un kļūst par pamatu turpmākajiem uzņēmējdarbības procesiem, tostarp izmaksu aprēķināšanai, plānošanai, aprites cikla novērtēšanai un integrācijai ar ERP un CAFM -sistēmām, kur procesu efektivitāte lielā mērā ir atkarīga no datu kvalitātes, kas nāk no projektēšanas nodaļām.

Tradicionālā pieeja atribūtu validācijai CAD- (BIM-) modeļos ietver manuālu validāciju (7.2-1. attēls), kas kļūst ilgs un dārgs process, ja modeļu apjoms ir liels. Nemot vērā mūsdienu būvniecības projektu apjomu un skaitu, kā arī to regulāros atjauninājumus, datu validēšanas un pārveidošanas process kļūst neilgtspējīgs un finansiāli neiespējams.

Ģenerāluzņēmējiem un projektu vadītājiem ir jāapstrādā liels projektu datu apjoms, tostarp vienas un tās pašas versijas un to pašu modeļu fragmenti. Datus saņem no projektēšanas organizācijām RVT, DWG, DGN, IFC, NWD un citos formātos (3.1-14. attēls), un ir regulāri jāpārbauda, vai tie atbilst nozares un korporatīvajiem standartiem.

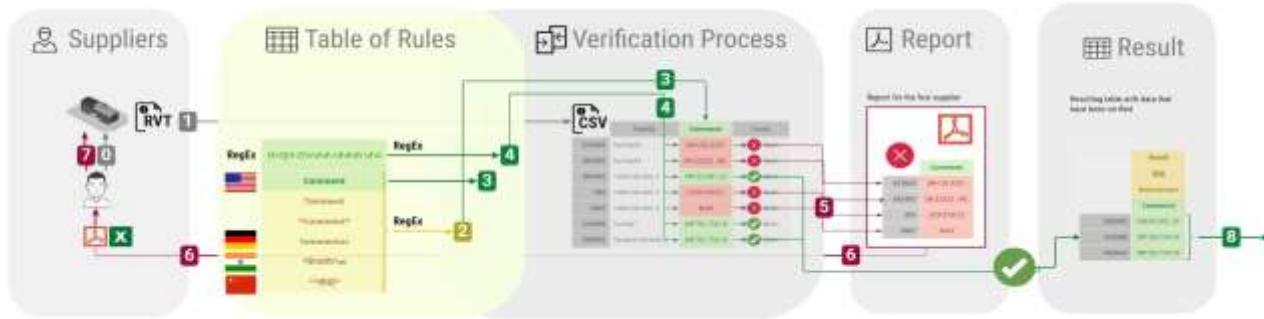
Atkarība no manuālām darbībām un specializētās programmatūras padara datu validēšanas procesu par šauru vietu darba plūsmās, kas saistītas ar datiem no uzņēmuma mēroga modeļiem. Automatizācija un strukturētu prasību izmantošana var novērst šo atkarību, ievērojami palielinot datu validēšanas ātrumu un uzticamību ( 7.3-7. attēls).



Attēls 7.3-7 Automatizācija palielina datu pārbaudes un apstrādes ātrumu, kas samazina darba izmaksas desmitiem reižu [140].

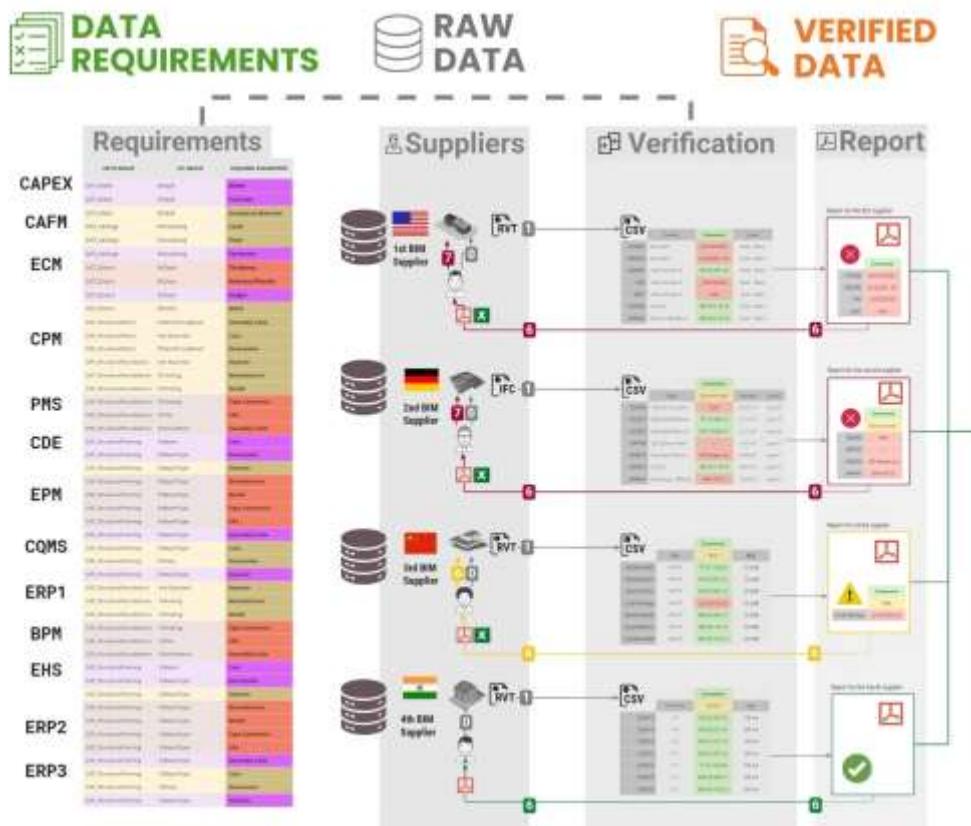
CAD datu validēšanas process ietver datu ieguvi (ETL posms Extract) no dažādiem slēgtiem (RVT, DWG, DGN, NWS u.c.) vai atvērtiem daļēji strukturētiem un parametriskiem formātiem (IFC, CPXML, USD.) vai atvērtos daļēji strukturētos un parametriskos formātos (IFC, CPXML, USD), kuros katram atribūtam un tā vērtībām var piemērot noteikumu tabulas (Transform posms), izmantojot regulārās izteiksmes RegEx ( 7.3-8. attēls), un šis process tika detalizēti aplūkots grāmatas ceturtajā daļā.

Klūdu ziņojuma izveide PDF formātā par un veiksmīgi validētiem ierakstiem jāpabeidz ar izvades datiem (Load step) strukturētos formātos, kuros nemitīgās vērā tikai validētas vienības, ko var izmantot turpmākajos procesos.



Attēls 7.3-8 Datu validācijas process no projekta datu sniedzējiem līdz gala ziņojumam, kas validēts, izmantojot regulārās izteiksmes.

Automatizējot datu validāciju no CAD sistēmām (BIM) ar strukturētām prasībām un jaunu datu plūsmu, kas tiek apstrādāti, izmantojot ETL cauruļvadus (7.3.3-9. attēls), tiek samazināta nepieciešamība validācijas procesā iesaistīties manuāli (katrs no validācijas un datu prasību procesiem ir apskatīts iepriekšējās nodalās).



Attēls 7.3-9 Datu validēšanas automatizēšana, izmantojot ETL, vienkāršo būvniecības projektu vadību, paātrinot procesus.

Tradicionāli darbuzņēmēju un CAD (BIM) specialistu sniegtu modeļu validācija var aizņemt no dažām

dienām līdz nedēļām. Tomēr, ieviešot automatizētus ETL procesus, šo laiku var samazināt līdz dažām minūtēm. Tipiskā situācijā būvuzņēmējs norāda: "*Modelis ir apstiprināts un atbilst prasībām.*" Ar šo apgalvojumu sākas ķedes verifikācija, lai pārbaudītu izpildītāja apgalvojumu par datu kvalitāti:

- ⌚ Projekta vadītājs - "Būvuzņēmējs apgalvo: "*Modelis ir pārbaudīts, viss ir kārtībā.*"."
- ⌚ Datu pārvaldnies - ielādes apstiprināšana:
  - Vienkāršs skripts programmā Pandas pārkāpumu atklāj dažu sekunžu laikā. Automatizācija novērš strīdus:
    - Kategorija: OST\_StructuralColumns, Parametrs: FireRating IS NULL.
    - Izveidot pārkāpumu ID sarakstu → eksportēt uz Excel/PDF.

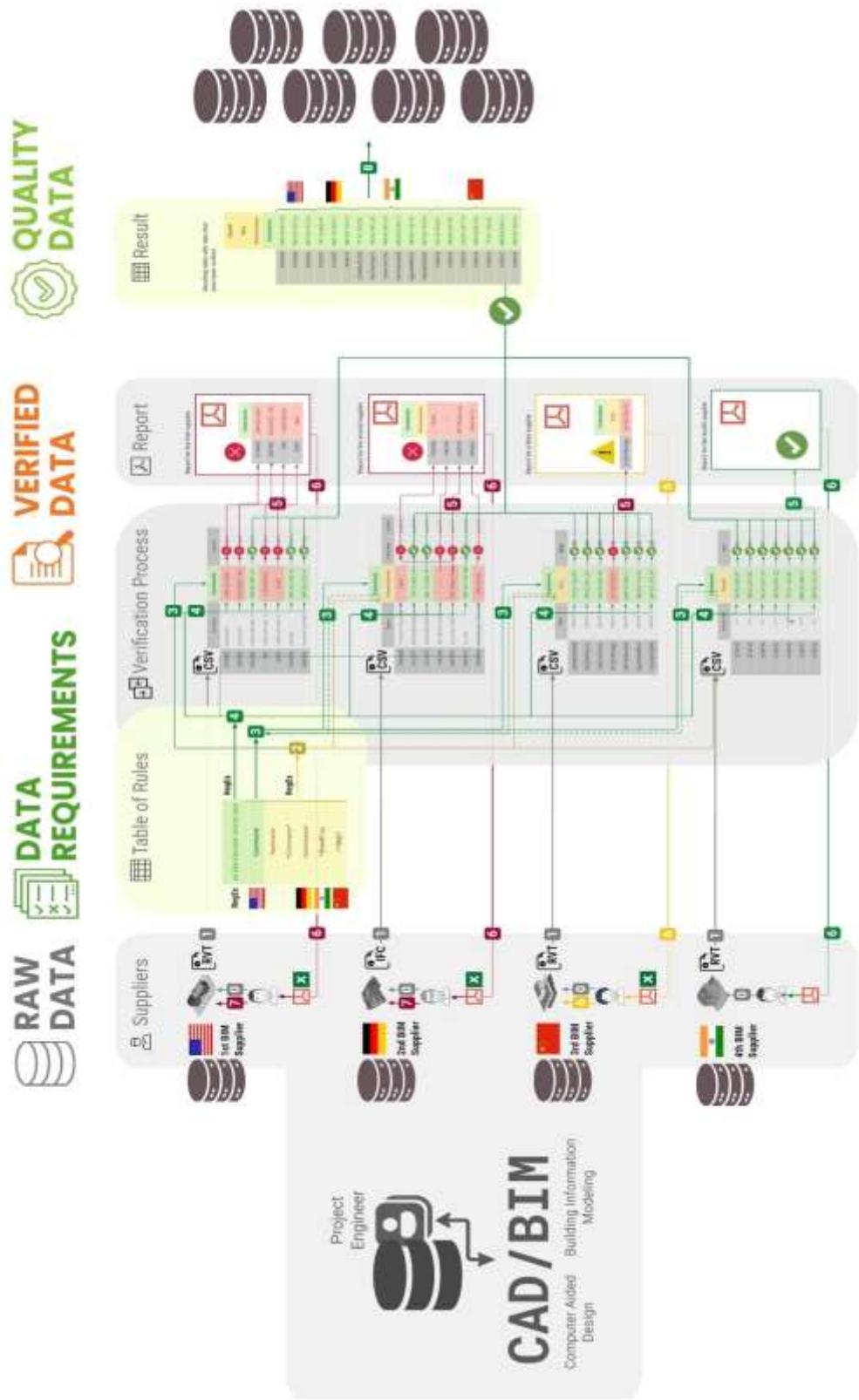
Vienkāršs skripts, kas darbojas ar programmu Pandas, pārkāpumu atklāj dažu sekunžu laikā:

```
df = model_data[model_data["Kategorija"] == "OST_StructuralColumns"] # Filtrēšana
issues = df[df["FireRating"].isnull()]] # Tukšas vērtības
issues[["ElementID"]].to_excel("fire_rating_issues.xlsx") # Eksporta ID
```

- ⌚ Datu pārvaldnies projekta vadītājam - "Pārbaudot, redzams, ka 18 slejās nav aizpildīts parametrs FireRating."
- ⌚ Projekta vadītājs darbuzņēmējam - "Modelis tiek atgriezts pārskatīšanai: parametrs FireRating ir obligāts, bez tā pieņemšana nav iespējama".

Rezultātā CAD modelis netiek validēts, automatizācija novērš strīdus, un darbuzņēmējs gandrīz uzreiz saņem strukturētu ziņojumu ar problemātisko elementu ID sarakstu. Šādā veidā validācijas process klūst pārredzams, atkārtojams un aizsargāts no cilvēka kļūdām (7.3-10. attēls).

Šī pieeja pārvērš datu validēšanas procesu par inženiertehnisku funkciju, nevis par manuālu kvalitātes kontroli. Tas ne tikai palielina produktivitāti, bet arī ļauj piemērot vienu un to pašu loģiku visiem uzņēmuma projektiem, nodrošinot visaptverošu digitālo transformāciju procesos no projektēšanas līdz operācijām.



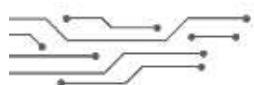
Attēls 7.3-10 Automatizējot elementu atribūtu pārbaudi, tiek novērsta cilvēka klūda un samazināta klūdu iespējamība.

Izmantojot automatizētus cauruļvadus ( 7.3-10. attēls), sistēmas lietotāji, kas sagaida kvalitatīvus datus no CAD- (BIM-) sistēmām, var uzreiz saņemt nepieciešamos izejas datus - tabulas, dokumentus, attēlus - un ātri integrēt tos savos darba uzdevumos.

Kontroles, apstrādes un analīzes automatizācija veicina izmaiņas būvniecības projektu vadības pieejā, jo īpaši dažādu sistēmu sadarbspēju, neizmantojot sarežģitas un dārgas modulārās patentētas sistēmas vai slēgtus piegādātāju risinājumus.

Lai gan jēdzieni un mārketinga akronīmi nāk un zūd, datu prasību apstiprināšanas procesi uz visiem laikiem paliks neatņemama biznesa procesu sastāvdaļa. Tā vietā, lai radītu arvien vairāk un vairāk specializētu formātu un standartu, būvniecības nozarei vajadzētu pievērsties rīkiem, kas jau ir pierādījuši savu efektivitāti citās nozarēs. Mūsdienās ir pieejamas jaudīgas datu apstrādes un procesu integrācijas automatizācijas platformas, kas ļauj uzņēmumiem ievērojami samazināt rutīnas operāciju veikšanai patēriņto laiku un līdz minimumam samazināt kļūdas, veicot Extract, Transform un Load.

Viens no populārākajiem risinājumu piemēriem ETL procesu automatizācijai un orķestrēšanai ir Apache Airflow, kas ļauj organizēt sarežģītus skaitlošanas procesus un pārvaldīt ETL cauruļvadus. Līdzās Airflow aktīvi tiek izmantoti arī citi līdzīgi risinājumi, piemēram, Apache NiFi datu maršrutēšanai un straumēšanai un n8n biznesa procesu automatizācijai.



## NODAĻA 7.4.

### ETL UN DARBA PLŪSMU ORKESTRĒŠANA: PRAKTISKI RISINĀJUMI

#### DAG un Apache Airflow: darba plūsmas automatizācija un orkestrācija

Apache Airflow ir bezmaksas atvērtā pirmkoda platforma, kas paredzēta darba plūsmu (ETL - konveijeru) automatizēšanai, organizēšanai un uzraudzībai.

Darbs ar lieliem datu apjomiem ir nepieciešams katru dienu:

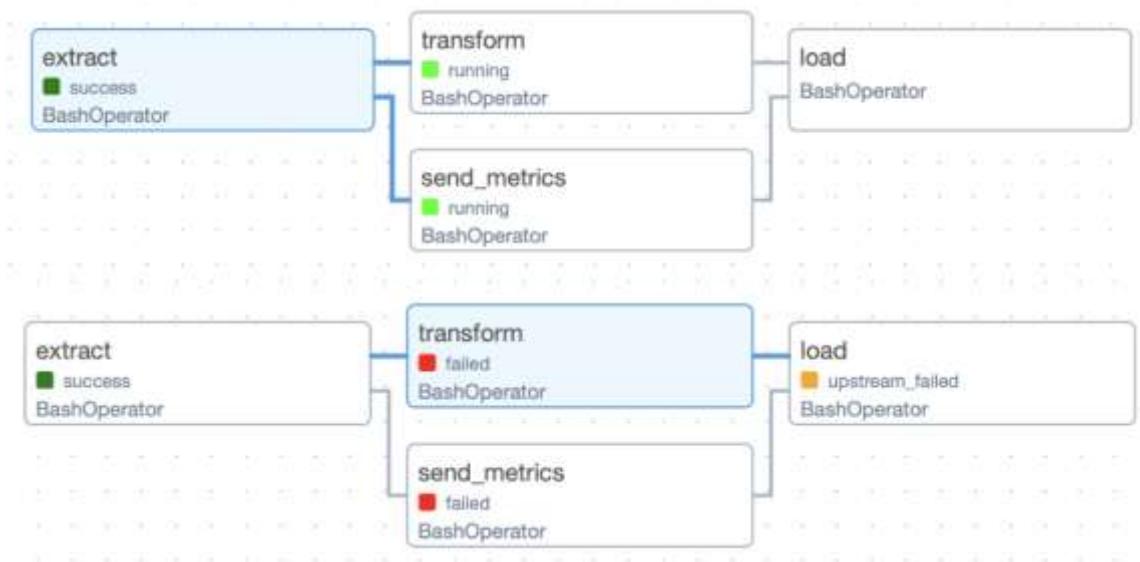
- Lejupielādējiet failus no dažādiem avotiem - Izraksts (piemēram, no piegādātājiem vai klientiem).
- pārveidot šos datus vajadzīgajā formātā - pārveidot (strukturēt, attīrīt un apstiprināt).
- Nosūtīt rezultātus pārbaudei un izveidot pārskatus - ielādēt (augšpielādēt vajadzīgajās sistēmās, dokumentos, datubāzēs vai vadības paneļos).

Šādu ETL procesu manuāla izpilde aizņem daudz laika un rada cilvēciskas kļūdas risku. Izmaiņas datu avotā vai kļūda kādā no posmiem var izraisīt kavēšanos un nepareizus rezultātus.

Automatizācijas rīki, piemēram, Apache Airflow, ļauj izveidot uzticamu ETL -konveijeru, samazināt kļūdas, saīsināt apstrādes laiku un nodrošināt, ka dati ir pareizi katrā posmā. Apache Airflow pamatā ir DAG (Directed Acyclic Graph) koncepcija - virzīts aciklisks grafiks, kurā katrs uzdevums (operators) ir saistīts ar citām atkarībām un tiek izpildīts stingri noteiktā secībā. DAG novērš ciklus, kas nodrošina loģisku un paredzamu uzdevumu izpildes struktūru.

Airflow rūpējas par orkestrēšanu - pārvalda atkarības starp uzdevumiem, kontrolē izpildes grafikus, seko līdzi statusam un automātiski reagē uz klūmēm. Šī pieeja samazina manuālo iejaukšanos un nodrošina visa procesa uzticamību.

**Uzdevumu orķestrators** ir rīks vai sistēma, kas paredzēta uzdevumu izpildes pārvaldībai un kontrolei sarežģītās skaitļošanas un informācijas vidēs. Tas atvieglo uzdevumu izpildes izvēršanas, automatizēšanas un pārvaldības procesu, lai uzlabotu veikspēju un optimizētu resursus.



**Attēls 7.4-1** Apache Airflow nodrošina lietotājam draudzīgu saskarni, kurā var vizualizēt DAG - ETL, apskatīt izpildes žurnālus, uzdevumu palaišanas statusu u. c.

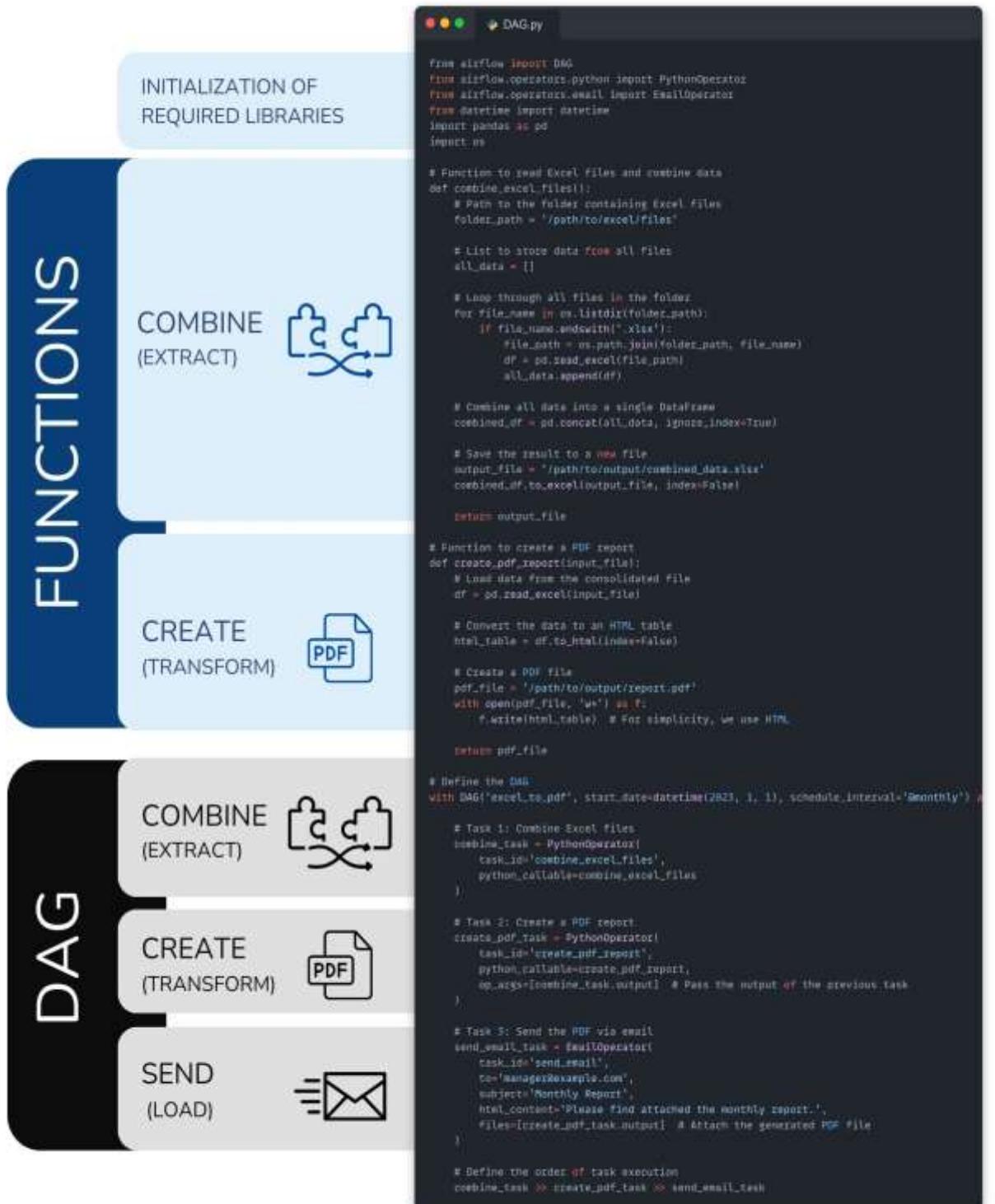
Airflow tiek plaši izmantots sadalītās skaitļošanas, datu apstrādes, ETL (Extract, Transform, Load) procesu pārvaldībai, uzdevumu plānošanai un automatizācijai, kā arī citiem datu scenārijiem. Pēc noklusējuma Apache Airflow kā datubāzi izmanto SQLite.

Vienkāršas DAG piemērs, kas līdzīgs ETL, sastāv no uzdevumiem - izvilkšana, pārveidošana un ielāde. Grafikā, kas tiek vadīts ar lietotāja saskarnes palīdzību (7.4-1. attēls), ir noteikta uzdevumu (koda fragmentu) izpildes secība: piemēram, vispirms tiek izpildīts extract (izvilkums), pēc tam transform (un nosūtīšana\_metriskas), un darbu pabeidz uzdevums load (ielāde). Kad visi uzdevumi ir izpildīti, datu ielādes process tiek uzskatīts par veiksmīgu.

## Apache Airflow: praktisks ETL automatizācijā

Apache Airflow tiek plaši izmantots, lai organizētu sarežģītus datu apstrādes procesus, ļaujot veidot elastīgus ETL-konveijerus. Apache Airflow var palaist vai nu ar tīmekļa saskarnes starpniecību, vai arī programmatiski, izmantojot Python kodu (7.4-2. attēls). Tīmekļa saskarnē (7.4-3. attēls) administratori un izstrādātāji var vizuāli uzraudzīt DAG, palaist uzdevumus un analizēt izpildes rezultātus.

Izmantojot DAG, varat noteikt skaidru uzdevumu secību, pārvaldīt atkarības starp tiem un automātiski reāgēt uz izmaiņām avota datos. Aplūkosim piemēru, kā izmantot Airflow, lai automatizētu pārskatu apstrādi (7.4-2. attēls).



Attēls 7.4-2 ETL -konveijera koncepcija datu apstrādei, izmantojot Apache Airflow.

Šajā piemērā ( 7.4-2. attēls) aplūkots DAG, kas veic galvenos uzdevumus ETL -konveijera ietvaros:

### ■ **Lasīt Excel - failus (Izraksts):**

- Sekvenciāla visu failu pārlūkošana dotajā direktorijā.
- Datu nolasīšana no katra faila, izmantojot pandas bibliotēku.
- Apvieno visus datus vienā DataFrame.

### ■ **Izveidot PDF -dokumentu (Transform):**

- Pārveidot apvienoto DataFrame HTML -tabulu.
- Saglabājiet tabulu kā PDF (demo versijā - caur HTML).

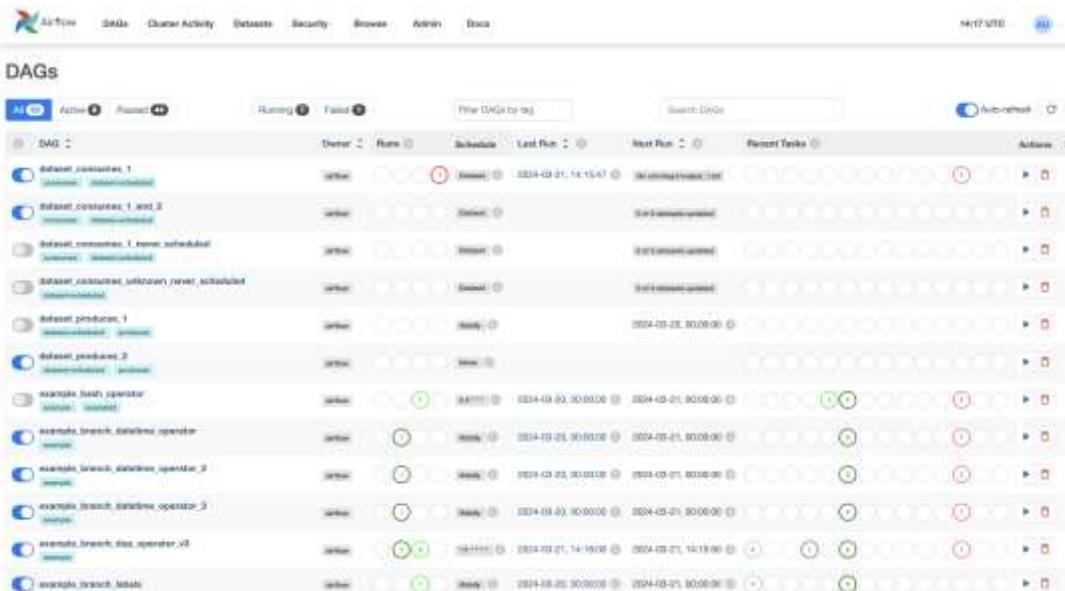
### ■ **Ziņojuma nosūtīšana pa e-pastu (ielāde):**

- Izmantojiet EmailOperator, lai nosūtītu PDF -dokumentu pa e-pastu.

### ■ **DAG konfigurēšana:**

- Uzdevumu secības noteikšana: datu iegūšana → pārskata ģenerēšana → nosūtīšana.
- Palaišanas grafika piešķiršana (@monthly - katra mēneša pirmā diena).

Automatizētajā ETL piemērā ( 7.4-2. attēls) parādīts, kā savākt datus no Excel failiem, izveidot PDF dokumentu un nosūtīt to pa e-pastu. Šis ir tikai viens no daudziem iespējamiem Airflow izmantošanas gadījumiem. Šo piemēru var pielāgot jebkuram konkrētam uzdevumam, lai vienkāršotu un automatizētu datu apstrādi.



**Attēls 7.4-3 Pārskats par visām DAG vidē ar informāciju par pēdējiem darbiem.**

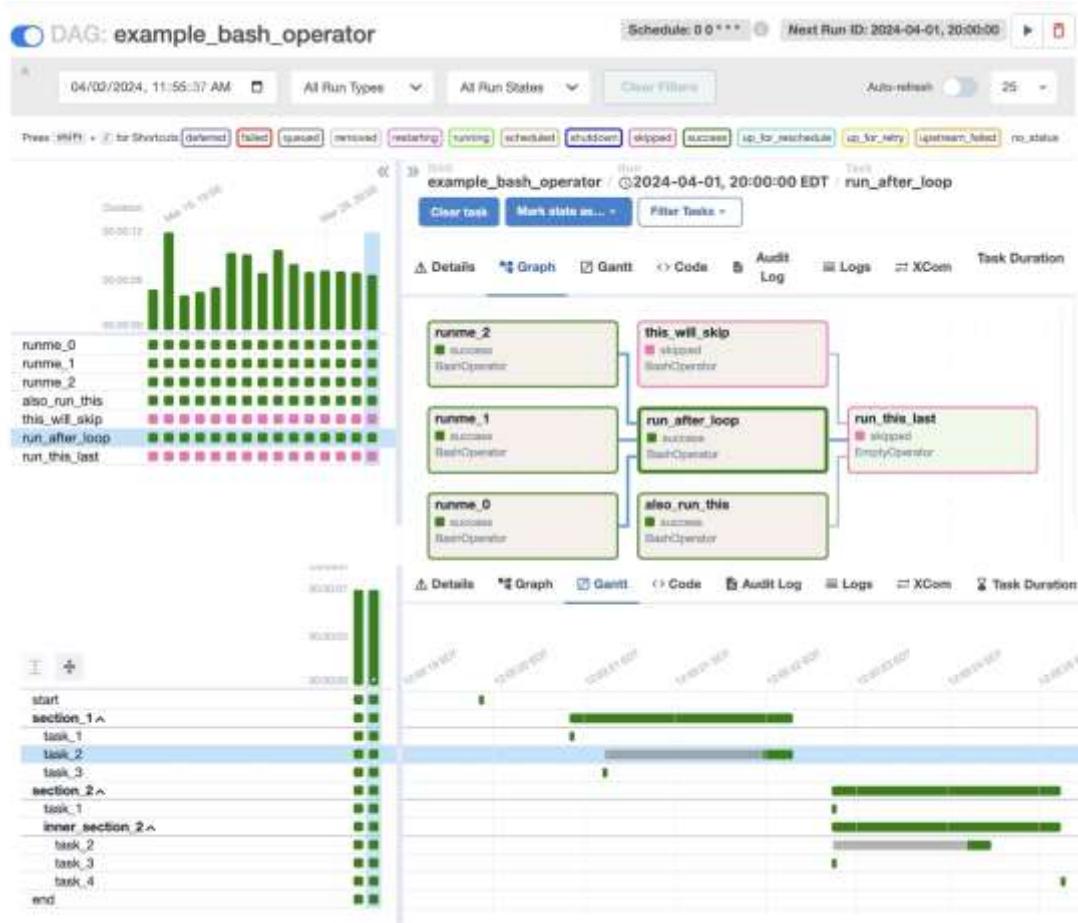
Apache Airflow tīmekļa saskarne ( 7.4-3. attēls) nodrošina visaptverošu vizuālo vidi datu darbplūsmu pārvaldībai. Tā attēlo DAG kā interaktīvus grafikus, kuros mezgli attēlo uzdevumus un malas - atkarības starp tiem, tādējādi ļaujot viegli izsekot sarežģītām datu plūsmām. Interfeiss ietver paneli ar informāciju par uzdevumu izpildes statusu, izpildes vēsturi, detalizētiem žurnāliem un veikspējas rādītājiem. Administratori var manuāli palaist uzdevumus, restartēt neveiksmīgas darbības, apturēt DAG un konfigurēt vides mainīgos, izmantojot intuitīvu lietotāja saskarni.

Šādu arhitektūru var papildināt ar datu validāciju, paziņojumiem par izpildes statusu, integrāciju ar

ārējiem API vai datubāzēm. Airflow ļauj elastīgi pielāgot DAG: pievienot jaunus uzdevumus, mainīt to secību, apvienot kēdes - tas padara to par efektīvu rīku sarežģītu datu apstrādes procesu automatizēšanai. Palaižot DAG Airflow tīmekļa saskarnē (7.4-3. attēls, 7.4-4. attēls), jūs varat uzraudzīt uzdevumu statusu. Sistēma izmanto krāsu indikāciju:

- Zaļa - uzdevums ir veiksmīgi izpildīts.
- Dzeltens - process ir procesā.
- Sarkana - kļūda uzdevuma izpildes laikā.

Ja rodas klūmes (piemēram, trūkst faila vai ir bojāta datu struktūra), sistēma automātiski sāk paziņojuma nosūtīšanu.



**Attēls 7.4-4 Apache Airflow ievērojami vienkāršo problēmu diagnostiku, procesu optimizāciju un komandas sadarbību sarežģītās datu apstrādes līnijās.**

Apache Airflow ir ērts, jo tas automatizē ikdienas uzdevumus, novēršot nepieciešamību tos veikt manuāli. Tā nodrošina uzticamību, uzraudzot procesu izpildi un sniedzot tūlītēju paziņojumu par kļūdām. Sistēmas elastīgums ļauj viegli pievienot jaunus uzdevumus vai mainīt esošos, pielāgojot darba plūsmas mainīgajām prasībām.

Papildus Apache Airflow ir līdzīgi rīki darbplūsmu orķestrēšanai. Piemēram, atvērtā koda un bezmaksas Prefect (7.3-5. attēls) piedāvā vienkāršāku sintaksi un labāk integrējas ar Python, Spotify izstrādātais Luigi nodrošina līdzīgu funkcionalitāti un labi darbojas ar lieliem datiem. Jāatzīmē arī Kronos un Dagster, kas piedāvā mūsdienīgas pieejas cauruļvadu veidošanai, galveno uzmanību pievēršot modularitātei un mērogojamībai. Uzdevumu orķestrēšanas rīka izvēle ir atkarīga no konkrētā projekta vajadzībām, taču tie visi palīdz automatizēt sarežģītus ETL datu procesus.

Īpaši jāatzīmē Apache NiFi - atvērtā pirmkoda platforma, kas paredzēta datu straumēšanai un maršrutēšanai. Atšķirībā no Airflow, kas koncentrējas uz sērijveida apstrādi un atkarību pārvaldību, NiFi koncentrējas uz reāllaika datu transformāciju un elastīgu maršrutēšanu starp sistēmām.

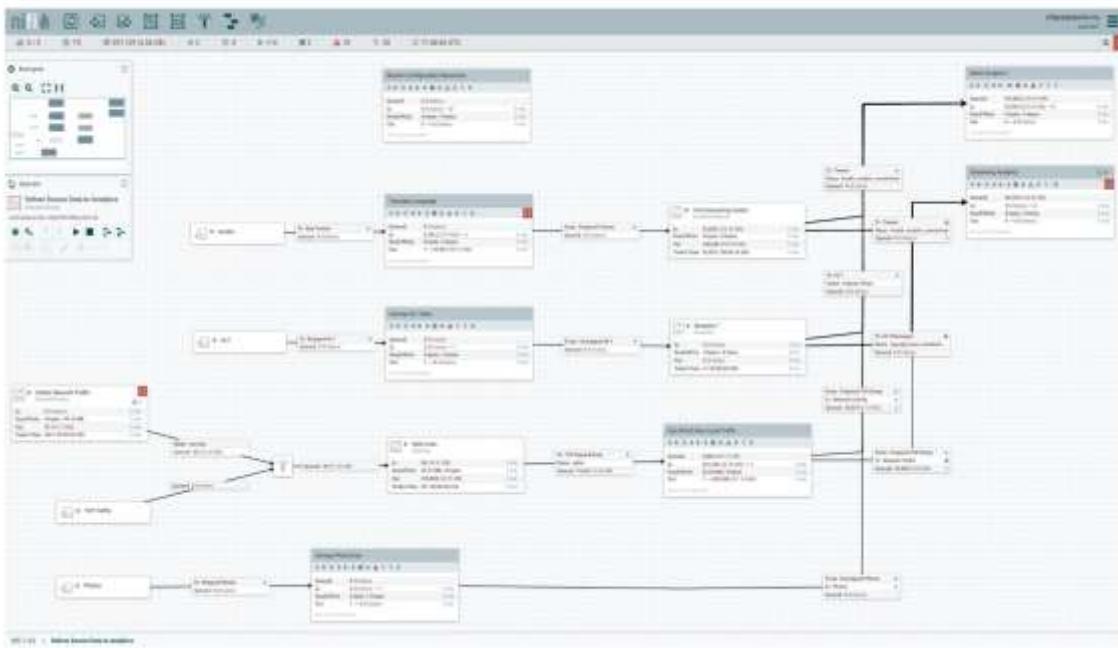
## Apache NiFi maršrutēšanai un datu konvertēšanai

Apache NiFi ir jaudīga atvērtā pirmkoda platforma, kas paredzēta datu plūsmu automatizēšanai starp dažādām sistēmām. Sākotnēji to 2006. gadā izstrādāja ASV Nacionālās drošības aģentūra (NSA) ar nosaukumu "Niagara Files" iekšējai lietošanai. Projekts 2014. gadā tika nodots kā atvērtā pirmavota projekts un nodots Apache Software Foundation, kļūstot par daļu no tā tehnoloģiju nodošanas iniciatīvām [141].

Apache NiFi ir paredzēts datu vākšanai, apstrādei un pārraidei reālajā laikā. Atšķirībā no Airflow, kas darbojas ar sērijveida uzdevumiem un prasa precīzi definētus grafikus, NiFi darbojas plūsmas apstrādes režīmā, ļaujot nepārtraukti pārsūtīt datus starp dažādiem pakalpojumiem.

Apache NiFi ir ideāli piemērots integrācijai ar IoT ierīcēm, būvniecības sensoriem, monitoringa sistēmām un, piemēram, CAD formātu straumēšanas validācijai serverī, kur var būt nepieciešama tūlītēja reakcija uz datu izmaiņām.

Ar iebūvētajiem filtrēšanas, transformēšanas un maršrutēšanas rīkiem NiFi ļauj standartizēt datus (Transform) pirms to pārsūtīšanas (Load) uz glabāšanas vai analītiskajām sistēmām. Viena no tās galvenajām priekšrocībām ir iebūvētais drošības atbalsts un piekļuves kontrole, kas padara to par uzticamu risinājumu konfidenciālas informācijas apstrādei.



Attēls 7.4-5 Datu plūsmas grafiskais attēlojums Apache NiFi saskarnē.

Apache NiFi efektīvi apstrādā reāllaika datu straumēšanas, filtrēšanas un maršrutēšanas uzdevumus. Tas ir ideāli piemērots tehniski intensīviem scenārijiem, kur svarīga ir stabila informācijas pārsūtīšana starp sistēmām un augsta caurlaides spēja.

Tomēr, ja galvenais mērķis ir integrēt dažādus pakalpojumus, automatizēt ikdienas darbības un ātri izveidot darba plūsmas bez padzījinātām programmēšanas zināšanām, ir nepieciešami risinājumi ar zemu sākuma slieksni un maksimālu elastību. Viens no šādiem rīkiem ir n8n - Low-Code /No-Code klases platforma, kas vērsta uz biznesa automatizāciju un vizuālu procesu orkestrēšanu.

## n8n Low-Code, No-Code procesu orkestrācija

**n8n** ir atvērtā koda platforma ar zemu kodu / bez koda automatizētu darba plūsmu izveidei, ko raksturo lietošanas ērtums, elastīgums un spēja ātri integrēties ar dažādiem ārējiem pakalpojumiem.

**No-Code** ir digitālo produkta radīšanas metode, nerakstot kodu. Visi procesa elementi - no loģikas līdz saskarnei - tiek realizēti tikai ar vizuālo rīku palīdzību. No-Code platformas ir paredzētas lietotājiem bez tehniskas sagatavotības un ļauj ātri izveidot automatizācjas, veidlapas, integrācijas un tīmekļa lietojumprogrammas. Piemērs: lietotājs izveido automātisku paziņojumu sūtīšanu vai Google Sheets integrāciju, izmantojot vilkšanas un nomešanas interfeisu bez programmēšanas zināšanām.

Izmantojot atvērtā koda un vietējās izvietošanas iespējas, n8n automatizācijas un ETL cauruļvadu izveides procesos nodrošina uzņēmumiem pilnīgu kontroli pār saviem datiem, vienlaikus garantējot

drošību un neatkarību no mākoņa pakalpojumu sniedzējiem.

Atšķirībā no Apache Airflow, kas ir orientēts uz skaitļošanas uzdevumiem ar stingru orkestrāciju un prasa Python zināšanas, n8n nodrošina vizuālo redaktoru, kas ļauj rakstīt skriptu, neprasot programmēšanas valodas zināšanas ( 7.4-6. attēls). Lai gan tā saskarne ļauj izveidot automatizētus procesus, nerakstot kodu (No-Code), sarežģītākos scenārijos lietotāji var pievienot savas JavaScript un Python -funkcijas, lai paplašinātu iespējas (Low-Code).

**Low-Code** ir programmatūras izstrādes pieeja, kurā lietojumprogrammas vai procesa pamata loģika tiek veidota, izmantojot grafisko saskarni un vizuālos elementus, un programmas kods tiek izmantots tikai funkcionalitātes pielāgošanai vai paplašināšanai. Low-Code platformas ļauj ievērojami paātrināt risinājumu izstrādi, iesaistot ne tikai programmētājus, bet arī biznesa lietotājus ar pamattehniskajām prasmēm. Piemērs: lietotājs var izveidot biznesa procesu no gataviem blokiem un vajadzības gadījumā pievienot savu skriptu JavaScript vai Python.

Lai gan n8n tiek pozicionēts kā platforma ar zemu sākuma slieksni, ir nepieciešamas pamata programmēšanas zināšanas, izpratne par tīmekļa tehnoloģijām un prasmes strādāt ar API. Sistēmas elastīgums ļauj to pielāgot visdažādākajiem uzdevumiem - no automatizētas datu apstrādes līdz integrācijai ar ziņojumapmaiņas, lietu interneta ierīcēm un mākoņpakkalpojumiem.

Galvenās n8n lietošanas funkcijas un priekšrocības:

- **Atvērtā pirmkoda** un vietējās izvietošanas iespējas nodrošina pilnīgu datu kontroli, atbilstību drošības prasībām un neatkarību no mākoņa pakalpojumu sniedzējiem.
- **Integrācija ar vairāk nekā 330 pakalpojumiem**, tostarp CRM, ERP, e-komerciju, mākonplatformām, ziņojumapmaiņām un datubāzēm.
- **Scenāriju elastība**: no vienkāršiem pazīnojumiem līdz sarežģītām kēdēm ar API apstrādi - pieprasījumiem, lēmumu loģiku un AI -pakkalpojumu pieslēgumu.
- **JavaScript un Python atbalsts**: lietotāji pēc vajadzības var ievietot pielāgotu kodu, paplašinot automatizācijas iespējas.
- **Intuitīva vizuālā saskarne**: ļauj ātri konfigurēt un vizualizēt visus procesa posmus.

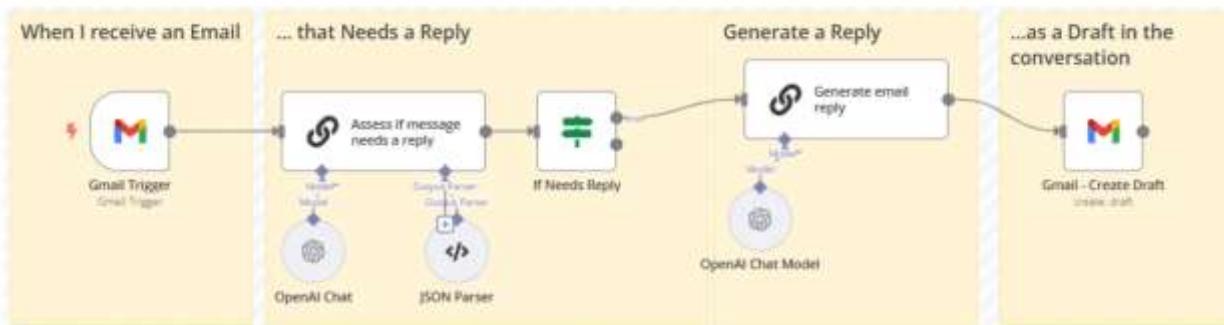
Low-Code klases platformas nodrošina rīkus digitālo risinājumu izveidei ar minimālu koda apjomu, tāpēc tās ir ideāli piemērotas komandām, kurām nav padziļinātu tehnisko zināšanu, bet ir nepieciešams automatizēt procesus.

Būvniecībā n8n var izmantot, lai automatizētu dažādus procesus, piemēram, integrētu ar projektu vadības sistēmām, pārbaudītu plūsmu, rakstītu gatavus pārskatus un vēstules, automātiski atjauninātu materiālu krājumu datus, nosūtītu komandām pazīnojumus par uzdevumu statusu un daudz ko citu. Pielāgots cauruļvads programmā n8n var vairākkārtīgi samazināt manuālo darbību skaitu, mazināt kļūdu iespējamību un paātrināt lēmumu pieņemšanu projekta īstenošanai.

Varat izvēlēties no gandrīz diviem tūkstošiem gatavu, bezmaksas un atvērtā pirmkoda n8n Pipeline, kas pieejams vietnē: [n8n.io/workflows](https://n8n.io/workflows), lai automatizētu gan būvniecības darba plūsmas, gan personīgos uzdevumus, samazinot rutīnas darbību skaitu.

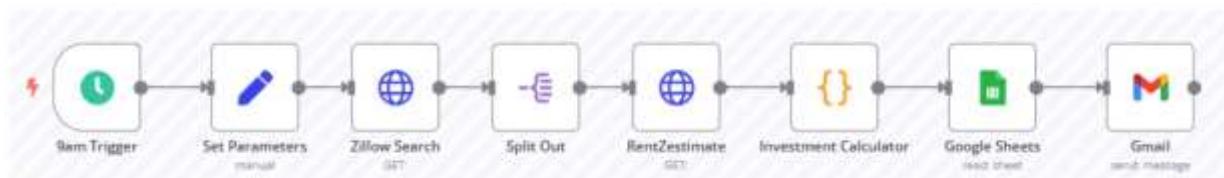
Nemiet piemēru vienu no gatavajām cauruļvadu veidnēm, kas bez maksas pieejama vietnē [n8n.io](https://n8n.io) [142], kas automātiski izveido atbilžu projektus pakalpojumā Gmail ( 7.4-6. attēls), palīdzot lietotājiem, kuri saņem lielu skaitu e-pasta vēstuļu vai kuriem ir grūtības sastādīt atbildes.

Šis n8n "Gmail AI Auto-Responder: Create draft responses to incoming emails" ( 7.4-6. attēls) veidne analizē ienākošos e-pasta ziņojumus, izmantojot LLM no ChatGPT, nosaka, vai ir nepieciešama atbilde, ģenerē projektu no ChatGPT un konvertē tekstu HTML un pievieno to ziņojumu ķēdē Gmail. Tas nenozīmē automātisku e-pasta vēstules nosūtīšanu, ja ujot jums manuāli redīģēt un apstiprināt atbildi. Iestatīšana aizņem aptuveni 10 minūtes, un tā ietver Gmail API OAuth konfigurāciju un OpenAI API integrāciju. Rezultāts ir ērts un bezmaksas risinājums ikdienas e-pasta saziņas automatizēšanai, nezaudējot kontroli pār e-pasta vēstuļu saturu.



Attēls 7.4-6 Automatizēts e-pasta atbildes ģenerēšanas process, izmantojot n8n.

Vēl viens automatizācijas piemērs, izmantojot n8n, ir izdevīgu piedāvājumu meklēšana nekustamā īpašuma tirgū [143]. N8n cauruļvads "Automatizēt ikdienas nekustamā īpašuma piedāvājumus, izmantojot Zillow API, Google Sheets un Gmail", katru dienu apkopo attiecīgos piedāvājumus, kas atbilst dotajiem kritērijiem, izmantojot Zillow API. Tā automātiski aprēķina galvenos ieguldījumu rādītājus (naudas līdzekļu atdevē, ikmēneša naudas plūsma, pirmā iemaksa), atjaunina Google Sheets un nosūta kopsavilkuma ziņojumu uz e-pastu (7.4-7. attēls), ja ujot investoriem ietaupīt laiku un ātri reaģēt uz labākajiem piedāvājumiem.



Attēls 7.4-7 Automatizēts nekustamā īpašuma investīciju pievilcības novērtēšanas process.

Tā elastīgums un paplašināmība padara n8n par vērtīgu rīku uzņēmumiem, kas vēlas digitāli pārveidoties un kļūt konkurētspējīgāki tirgū, izmantojot salīdzinoši vienkāršus un bezmaksas atvērtā

koda rīkus.

Tādus rīkus kā Apache NiFi, Airflow un n8n var uzskatīt par trīs datu apstrādes slāņiem (7.4.4-8. attēls). NiFi pārvalda datu plūsmu, nodrošinot to piegādi un pārveidošanu, Airflow organizē uzdevumu izpildi, apkopojot datus apstrādes plūsmās, bet n8n automatizē integrāciju ar ārējiem pakalpojumiem un pārvalda biznesa loģiku.



	The main task	Approach
<b>Apache NiFi</b>	Streaming and data transformation	Real-time stream processing
<b>Apache Airflow</b>	Task orchestration, ETL pipelines	Batch planning, DAG processes
<b>n8n</b>	Integration, automation of business logic	Low-code visual orchestration

**Attēls 7.4-8 Apache Airflow, Apache NiFi un n8n var uzskatīt par trim savstarpēji papildinošiem mūsdienu datu pārvaldības arhitektūras slāņiem.**

Kopā šie bezmaksas un atvērtā koda rīki potenciāli veido efektīvas datu un procesu pārvaldības ekosistēmas piemēru būvniecības nozarē, laujot uzņēmumiem izmantot informāciju lēmumu pieņemšanai un procesu automatizācijai.

## Nākamie soļi: pāreja no manuālām operācijām uz analītikā balstītiem risinājumiem

Mūsdienu būvniecības uzņēmumi darbojas augstas nenoteiktības apstākjos: mainīgas materiālu cenas, kavētas piegādes, darbaspēka trūkums un saspringti projektu termiņi. Analītisko paneļu, ETL - konveijeru un BI sistēmu izmantošana palīdz uzņēmumiem ātri identificēt problemātiskās jomas, novērtēt resursu efektivitāti un paredzēt izmaiņas, pirms tās rada finansiālus zaudējumus.

Apkopojoš šo daļu, ir vērts uzsvērt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotās tehnoloģijas ikdienas uzdevumos:

- Datu vizualizāciju un analītisko paneļu ieviešana
  - Apgūt informācijas paneļu izveides procesu, lai uzraudzītu galvenos darbības rādītājus (KPI).
  - Izmantojiet datu vizualizācijas rīkus (Power BI, Tableau, Matplotlib, Plotly).
- Datu apstrādes automatizēšana, izmantojot ETL -procesus.
  - Iestatiet automātisku datu vākšanu no dažādiem avotiem (dokumentācija, tabulas,

CAD), izmantojot ETL procesus.

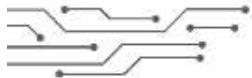
- organizēt datu pārveidošanu (piemēram, regulārās izteiksmes pārbaudi vai aprēķinus), izmantojot Python skriptus.
- Izmēģiniet iestatīt automātisku PDF (vai DOC) pārskatu sagatavošanu, izmantojot FPDF bibliotēku, izmantojot datus no Excel failiem vai iegūstot informāciju no citiem PDF dokumentiem.

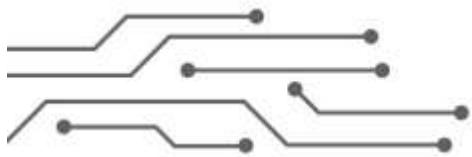
#### ■ Valodas modeļu (LLM) izmantošana automatizācijai

- Lielo valodas modeļu (LLM) izmantošana, lai ģenerētu kodu, kas palīdz iegūt un analizēt datus no nestukturētiem dokumentiem.
- Iepazīstieties ar n8n automatizācijas rīku un izpētiet gatavās veidnes un gadījumu izpētes viņu vietnē. Noskaidrojet, kurus jūsu darba procesus var pilnībā automatizēt, izmantojot metodi "No-Code/Low-Code".

Analītiska pieeja datu un procesu automatizācijai ne tikai samazina rutīnas darbību veikšanai patērēto laiku, bet arī uzlabo lēmumu pieņemšanas kvalitāti. Uzņēmumi, kas ievieš vizuālās analīzes rīkus un ETL -konveijerus, iegūst iespēju ātri reaģēt uz izmaiņām

Biznesa procesu automatizēšana, izmantojot tādus rīkus kā n8n, Airflow un NiFi, ir tikai pirmais solis ceļā uz digitālo briedumu. Nākamais solis ir kvalitatīva to datu glabāšana un pārvaldība, kas ir automatizācijas pamatā. Astotajā daļā mēs padziļināti aplūkosim, kā būvniecības uzņēmumi var izveidot ilgtspējīgu datu glabāšanas arhitektūru, pārejot no dokumentu un dažāda formāta failu haosa uz centralizētām glabāšanas un analīzes platformām.





## VIII DAĻA

# DATU GLABĀŠANA UN PĀRVALDĪBA BŪVΝIECĪBĀ

8. daļā aplūkotas modernas datu glabāšanas un pārvaldības tehnoloģijas būvniecības nozarē. Tajā analizēti efektīvi formāti lielu informācijas apjomu apstrādei - no vienkāršiem CSV un XLSX līdz jaudīgākiem Apache Parquet un ORC, detalizēti salīdzinot to iespējas un ierobežojumus. Apskatīti datu noliktavu (DWH), datu ezeru) un to hibrīdrisinājumu (Data Lakehouse), kā arī datu pārvaldības) un datu minimālisma) principi. Šīki aplūkotas datu purva) problēmas un stratēģijas, kā novērst haosu informācijas sistēmās. Tieks izklāstītas jaunas pieejas darbam ar datiem, tostarp vektoru datubāzes un to pielietojums konstruēšanā, izmantojot Bounding Box koncepciju. Šajā daļā ir aplūkotas arī DataOps un VectorOps metodoloģijas kā jauni datu darba plūsmu organizēšanas standarti.

## NODAĻA 8.1.

# DATU INFRASTRUKTŪRA: NO GLABĀŠANAS FORMĀTIEM LĪDZ DIGITĀLAJĀM KRĀTUVĒM

### Datu atomi: efektīvas informācijas pārvaldības pamats

Viss Visumā sastāv no mazākajiem būvmateriāliem - atomiem un molekulām, un laika gaitā visas dzīvās un nedzīvās lietas neizbēgami atgriežas šajā sākotnējā stāvoklī. Dabā šis process notiek ar pārsteidzošu ātrumu, ko mēs cenšamies pārnest uz cilvēka kontrolētiem procesiem.

Mežā visi dzīvie organismi galu galā pārvēršas barības vielā, kas kalpo par pamatu jauniem augiem. Savukārt šie augi kļūst par barību jaunām dzīvām būtnēm, kas sastāv no tiem pašiem atomiem, kuri pirms miljoniem gadu radīja Visumu.

Arī biznesa pasaule ir svarīgi sadalīt sarežģītas, daudzslāņainas struktūras to pamatvienībās, kas ir minimāli apstrādātas - līdzīgi kā atomi un molekulas dabā. Tas ļauj efektīvi uzglabāt un pārvaldīt datu atomus, pārvēršot tos par bagātīgu, auglīgu pamatu, kas kļūst par galveno resursu analītikas un lēmumu kvalitātes izaugsmei.



Attēls 8.1-1 Analīze un lēmumu pieņemšana balstās uz atkārtoti izmantotiem datiem, kas reiz jau ir apstrādāti un saglabāti.

Mūzikas skaņdarbus veido notis, kas apvienojas, veidojot sarežģītus skaņdarbus, savukārt vārdi tiek veidoti no primitīvas vienības - burta un skaņas. Vai tā būtu daba, zinātne, ekonomika, māksla vai

tehnoloģijas, pasaule uzrāda ievērojamu vienotību un harmoniju, cenšoties iznīcināt, strukturēt, ciklizēt un radīt. Līdzīgi arī procesi izmaksu aprēķina sistēmās tiek sadalīti sīkās strukturētās vienībās - resursu vienībās - izmaksu aprēķina un grafiku līmenī. Šīs vienības, tāpat kā piezīmes, pēc tam tiek izmantotas, lai veidotu sarežģītākus aprēķinus un grafikus. To pašu principu izmanto datorizētās projektēšanas sistēmas, kurās sarežģīti arhitektūras un inženier Tehniskie projekti tiek veidoti no pamatelementiem - atsevišķiem elementiem un bibliotēkas komponentiem, no kuriem tiek izveidots pilns sarežģītas ēkas vai būves projekta 3D -modelis.

Dabai un zinātnei raksturīgais cikliskuma un struktūras jēdziens atspogulojas arī mūsdienu datu pasaulē. Tāpat kā dabā visas dzīvās būtnes atgriežas pie atomiem un molekulām, tā arī mūsdienu datu apstrādes rīku pasaulē informācijai ir tendence atgriezties pie tās primitīvākās formas.

Mazākie elementi ar to galīgo nedalāmību ir uzņēmējdarbības procesu pamatelementi. Jau no paša sākuma ir svarīgi rūpīgi apsvērt, kā apkopot, strukturēt (sadalīt atomos) un uzglabāt šos sīkos pamatelementus no dažādiem avotiem. Datu organizēšana un glabāšana nav tikai to sadalīšana sastāvdaļas. Tikpat svarīgi ir nodrošināt to integrēšanu un uzglabāšanu strukturētā veidā, lai datus varētu viegli iegūt, analizēt un izmantot lēmumu pieņemšanai, kad vien tie ir nepieciešami.

Lai efektīvi apstrādātu informāciju, ir rūpīgi jāizvēlas datu glabāšanas formāts un metodes - tāpat kā augsne ir jāsagatavo koku augšanai. Datu krātuves ir jāorganizē tā, lai nodrošinātu augstu informācijas kvalitāti un atbilstību, izslēdzot liekus vai nebūtiskus datus. Jo labāk šī "informācijas augsne" ir strukturēta, jo ātrāk un precīzāk lietotāji var atrast vajadzīgos datus un atrisināt analītiskas problēmas.

## **Informācijas glabāšana: faili vai dati**

Datu noliktavas ļauj uzņēumiem apkopot un apvienot informāciju no dažādām sistēmām, izveidojot vienu centru turpmākai analīzei. Apkopotie vēsturiskie dati ļauj ne tikai padziļināti analizēt procesus, bet arī identificēt likumsakarības, kas var ietekmēt uzņēmuma darbību.

Pieņemsim, ka uzņēmums vienlaikus strādā pie vairākiem projektiem. Inženieris vēlas saprast, cik daudz betona ir ieliets un cik daudz vēl jāpērk. Izmantojot tradicionālo pieeju, viņam būtu manuāli jāmeklē serverī un jāatver vairākas tāmju tabulas, jāsalīdzina tās ar izpildīto darbu apliecinājumiem un jāpārbauda pašreizējie krājumu atlikumi. Tas aizņem vairākas stundas vai pat dienas. Pat izmantojot ETL procesus un automātiskos skriptus, uzdevums joprojām ir daļēji manuāls: inženierim joprojām ir manuāli jānorāda ceļš līdz mapēm vai konkrētiem failiem serverī. Tas samazina kopējo automatizācijas efektu, jo joprojām aizņem vērtīgu darba laiku.

Pārejot uz datu pārvaldību, tā vietā, lai strādātu ar servera failu sistēmu, inženieris iegūst piekļuvi vienotai glabāšanas struktūrai, kurā informācija tiek atjaunināta reālajā laikā. Ar vienu vienīgu vaicājumu - koda, SQL vai pat LLM izsaukuma veidā - var uzreiz iegūt precīzus datus par pašreizējiem atlikumiem, veikto darbu apjomiem un gaidāmajām piegādēm, ja dati ir iepriekš sagatavoti un apvienoti datu krātuvē, kur nav nepieciešams klejot pa mapēm, atvērt desmitiem failu un manuāli salīdzināt vērtības.

Ilgū laiku būvniecības uzņēmumi izmantoja PDF -dokumentus, DWG - rasējumus, RVT - modeļus un simtiem un tūkstošiem Excel - tabulu un citus atšķirīgus formātus, kas tiek glabāti noteiktās mapēs uzņēmuma serveros, apgrūtinot informācijas meklēšanu, pārbaudi un analīzi. Rezultātā faili, kas paliek pēc projektu pabeigšanas, visbiežāk tiek pārvietoti atpakaļ uz serveri arhīva glabāšanas mapēs, kuras turpmāk praktiski netiek izmantotas. Šāda tradicionālā uz failiem balstīta datu glabāšana zaudē aktualitāti, pieaugot datu plūsmai, jo tā ir neaizsargāta pret cilvēka klūdām.

**Datne** ir tikai atsevišķs konteiners, kurā tiek glabāti dati. Faili ir radīti cilvēkiem, nevis sistēmām, tāpēc tos ir nepieciešams atvērt, nolasīt un interpretēt manuāli. Piemēram, Excel -tabula, PDF -dokuments vai CAD -zīmējums, kas ir īpaši jāatver konkrētā rīkā, lai piekļūtu vajadzīgajai informācijai. Bez strukturētas izguves un apstrādes tajā esošā informācija paliek neizmantota.

Savukārt **dati** ir mašīnlasāma informācija, kas tiek automātiski savienota, atjaunināta un analizēta. Atsevišķā datu krātuvē (piemēram, datu bāzē, DWH vai datu ezerā) informācija tiek attēlotā tabulu, ierakstu un attiecību veidā. Tas nodrošina vienotu glabāšanu, automatizētus pieprasījumus, vērtību analīzi un pārskatu sniegšanu reāllaikā.

Datu izmantošana datņu vietā ( 8.1-1. attēls ) ļauj novērst manuālās meklēšanas procesu un unificēt apstrādi. Uzņēmumi, kas jau ir ieviesuši šādu pieejumu, iegūst konkurences priekšrocības, jo informācijai ir ātri pieejama un to var ātri integrēt uzņēmējdarbības procesos.

Pāreja no failu izmantošanas uz datu izmantošanu ir neizbēgama pārmaiņa, kas noteiks būvniecības nozares nākotni.

Ikvienam būvniecības nozares uzņēmumam būs jāizvēlas: vai turpināt glabāt informāciju atšķirīgos failos un silosos, kas jālasa cilvēkiem, izmantojot īpašas programmas, vai pārveidot to pirmajos apstrādes posmos strukturētos datos, izveidojot vienotu integrētu digitālo pamatu automatizētai projektu pārvaldībai.

Files	Data
Stored in isolated structures	Open and accessible information flows
Stored in folders with limited access for a few employees	Stored in centralized databases and accessible across multiple processes
Require manual search, opening, and analysis	Can be automatically processed and analyzed in real time
Difficult to integrate between systems	Easily integrated and matched for analytics

Attēls 8.1-1 Informācijas plūsmas attīstība: no izolētām datnēm līdz integrētiem datiem.

Līdz ar informācijas sprādzeniem tradicionālās failu glabāšanas un apstrādes metodes kļūst arvien mazāk efektīvas. Būvniecības nozarē, tāpat kā citās nozarēs, vairs nepietiek paļauties uz atšķirīgām datņu mapēm ar dažādiem datņu formātiem vai nesaistītām datubāzēm.

Uzņēmumi, kas vēlas saglabāt konkurētspēju digitālajā laikmetā, neizbēgami pāries uz integrētām digitālajām platformām, izmantos lielo datu tehnoloģijas un automatizētas analītikas sistēmas.

Pāreja no uz datnēm balstītās uzglabāšanas uz datu plūsmām prasīs informācijas pārvaldības pieeju pārskatīšanu un apzinātu tādu formātu izvēli, kas piemēroti turpmākai integrācijai centralizētās krātuvēs. Šī izvēle noteiks, cik efektīvi var apstrādāt datus, cik ātri tiem var piekļūt un cik viegli tos var integrēt uzņēmuma digitālajos procesos.

## Lielo datu glabāšana: populāru formātu un to efektivitātes analīze

Uzglabāšanas formātiem ir būtiska nozīme analītikas infrastruktūras mērogojamībā, uzticamībā un veiktspējā. Datu analīzei un apstrādei, piemēram, filtrēšanai, grupēšanai un apkopošanai, mūsu piemēros tika izmantots Pandas DataFrame - populāra struktūra darbam ar datiem RAM.

Tomēr Pandas DataFrame nav sava glabāšanas formāta, tāpēc pēc apstrādes pabeigšanas dati tiek eksportēti kādā no ārējiem formātiem - visbiežāk CSV vai XLSX. Šos tabulu formātus ir viegli apmainīt un tie ir saderīgi ar lielāko daļu ārējo sistēmu, taču tiem ir vairāki ierobežojumi: zema uzglabāšanas efektivitāte, saspiešanas trūkums un vājš versiju atbalsts:

- **CSV** (Comma-Separated Values): vienkāršs teksta formāts, ko plaši atbalsta dažādas platformas un rīki. Tas ir viegli lietojams, bet neatbalsta sarežģītus datu tipus un saspiešanu.
- **XLSX** (Excel Open XML Spreadsheet): Microsoft Excel faila formāts, kas atbalsta tādas sarežģītās funkcijas kā formulas, diagrammas un stilu. Lai gan tas ir noderīgs manuālai datu analīzei un vizualizācijai, tas nav optimizēts liela apjoma datu apstrādei.

Papildus populārajiem tabulārajiem XLSX un CSV, ir vairāki populāri formāti efektīvai strukturētu datu glabāšanai ( 8.1-2. attēls), un katram no tiem ir unikālas priekšrocības atkarībā no konkrētām datu glabāšanas un analīzes prasībām:

- **Apache Parquet:** kolonnveida datu glabāšanas datņu formāts, kas optimizēts izmantošanai datu analīzes sistēmās. Tas piedāvā efektīvas datu saspiešanas un kodēšanas shēmas, padarot to ideāli piemērotu sarežģītām datu struktūrām un lielo datu apstrādei.
- **Apache ORC (Optimised Row Columnar):** līdzīgi kā Parquet, ORC nodrošina augstu saspiešanas pakāpi un efektīvu datu glabāšanu. Tas ir optimizēts smagām nolasīšanas operācijām un ir labi piemērots datu ezeru glabāšanai.
- **JSON (JavaScript Object Notation):** lai gan JSON nav tik efektīvs datu glabāšanas ziņā salīdzinājumā ar binārajiem formātiem, piemēram, Parquet vai ORC, tas ir ļoti pieejams un viegli lietojams, tāpēc ir ideāli piemērots skriptiem, kur svarīga ir lasāmība un tīmekļa saderība.
- **Feather:** ātrs, viegls un viegli lietojams uz analītiku orientēts bināro kolonnveida datu glabāšanas formāts. Tas ir izstrādāts efektīvai datu pārsūtīšanai starp Python (Pandas) un R, padarot to par lielisku izvēli projektiem, kas saistīti ar šīm programmēšanas vidēm.
- **HDF5 (Hierarchiskais datu formāts, 5. versija):** paredzēts liela datu apjoma glabāšanai un organizēšanai. Tas atbalsta plašu datu tipu klāstu un ir labi piemērots darbam ar sarežģītām datu kopām. HDF5 ir īpaši populārs zinātniskajos skaitlošanas procesos, jo tas ļauj efektīvi uzglabāt un piekļūt lielām datu kopām.

	XLSX	CSV	Apache Parquet	HDF5	Pandas DataFrame
Storage	Tabular	Tabular	Columnar	Hierarchical	Tabular
Usage	Office tasks, data presentation	Simple data exchange	Big data, analytics	Scientific data, large volumes	Data analysis, manipulation
Compression	Built-in	None	High	Built-in	None (in-memory)
Performance	Low	Medium	High	High	High (memory dependent)
Complexity	High (formatting, styles)	Low	Medium	Medium	Low
Data Type Support	Limited	Very limited	Extended	Extended	Extended
Scalability	Low	Low	High	High	Medium (memory limited)



Attēls 8.1-2 Datu formātu salīdzinājums, kurā redzamas galvenās atšķirības glabāšanas un apstrādes aspektos.

Lai veiktu ETL procesa Load posmā izmantoto formātu salīdzinošo analīzi, tika izveidota tabula, kurā parādīti failu izmēri un nolasīšanas laiki ( 8.1-3. attēls). Pētījumā tika izmantoti faili ar identiskiem datiem: tabulā bija 10 000 rindu un 10 slejas, kas aizpildītas ar izlases vērtībām.

Pētījumā ir iekļauti šādi glabāšanas formāti: CSV, Parquet, XLSX un HDF5, kā arī to saspiecēs versijas ZIP arhīvos. Neapstrādātie dati tika ģenerēti, izmantojot NumPy bibliotēku, un attēloti kā Pandas DataFrame struktūra. Testēšanas process sastāvēja no šādiem posmiem:

- Failu saglabāšana: datu kopa tiek saglabāta četros dažādos formātos: CSV, Parquet, XLSX un HDF5. Katram formātam ir unikālas datu glabāšanas īpatnības, kas ietekmē faila lielumu un nolasīšanas ātrumu.
- ZIP faila saspiešana: lai analizētu standarta saspiešanas efektivitāti, katrs fails tika saspiests ZIP arhīvā.
- Failu nolasīšana (ETL - Load): nolasīšanas laiks tika mērīts katram failam pēc tā iziepkakošanas no ZIP. Tas ļauj novērtēt datu piekļuves ātrumu pēc izvilkšanas no arhīva.

Iz svarīgi atzīmēt, ka Pandas DataFrame netika tieši izmantots lieluma vai nolasīšanas laika analīzēs, jo tas nav atsevišķs glabāšanas formāts. Tā kalpoja tikai kā starposma struktūra datu ģenerēšanai un turpmākai saglabāšanai dažādos formātos.



Attēls 8.1-3 Uzglabāšanas formātu salīdzinājums pēc izmēra un nolasīšanas ātruma.

CSV un HDF5 faili uzrāda ( 8.1-3. attēls) augstu saspiešanas efektivitāti, ievērojami samazinot to izmēru, kad tie ir iepakoti ZIP failā, kas var būt īpaši noderīgi scenārijos, kuros nepieciešama uzglabāšanas optimizācija. No otras puses, XLSX faili ir praktiski nesaspiežami, un to izmērs ZIP formātā paliek salīdzināms ar oriģināla izmēru, tāpēc tie nav tik piemēroti izmantošanai lielos datu

apjomos vai vidēs, kur svarīgs ir datu piekļuves ātrums. Turklāt XLSX lasīšanas laiks ir ievērojami ilgāks salīdzinājumā ar citiem formātiem, tāpēc tas ir mazāk piemērots ātrai datu lasīšanai. Apache Parquet ir demonstrējis augstu veiktspēju analītiskiem uzdevumiem un lieliem datu apjomiem, pateicoties tā kolonnveida struktūrai.

## Uzglabāšanas optimizēšana ar Apache Parquet

Viens no populārākajiem lielo datu glabāšanas un apstrādes formātiem ir Apache Parquet. Šis formāts ir īpaši izstrādāts kolonnveida glabāšanai (līdzīgi kā Pandas), kas ļauj ievērojami samazināt atmiņas ietilpību un palielināt analītisko vaicājumu ātrumu. Atšķirībā no tradicionālajiem formātiem, piemēram, CSV un XLSX, Parquet atbalsta dabisko saspiešanu un ir optimizēts lielo datu sistēmām, tostarp Spark, Hadoop un mākoņglabāšanai.

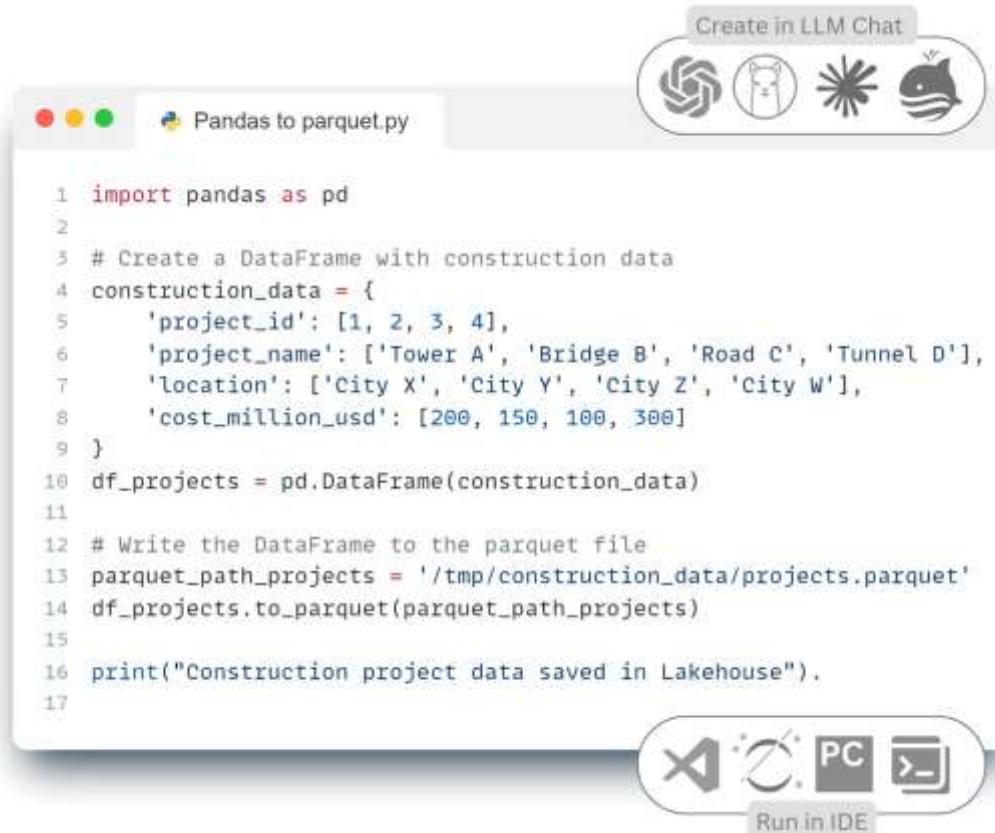
Parquet galvenās funkcijas ietver datu saspiešanas un kodēšanas atbalstu, kas ievērojami samazina datu glabāšanas apjomu un paātrina datu nolasīšanas operācijas, darbojoties tieši ar vēlamajām kolonnām, nevis visām datu rindām.

Lai uzskatāmi parādītu, cik viegli ir iegūt vajadzīgo kodu, lai konvertētu datus uz Apache Parquet, izmantosim LLM.

- 💡 Nosūtīt teksta pieprasījumu LLM tērzēšanai (CHATGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN):

Uzrakstiet kodu, lai saglabātu datus no Pandas DataFrame uz Apache Parquet. ↵

❓ LLM atbilde :



```

1 import pandas as pd
2
3 # Create a DataFrame with construction data
4 construction_data = {
5     'project_id': [1, 2, 3, 4],
6     'project_name': ['Tower A', 'Bridge B', 'Road C', 'Tunnel D'],
7     'location': ['City X', 'City Y', 'City Z', 'City W'],
8     'cost_million_usd': [200, 150, 100, 300]
9 }
10 df_projects = pd.DataFrame(construction_data)
11
12 # Write the DataFrame to the parquet file
13 parquet_path_projects = '/tmp/construction_data/projects.parquet'
14 df_projects.to_parquet(parquet_path_projects)
15
16 print("Construction project data saved in Lakehouse").
17

```

**Attēls 8.1-4** Datu kadru datu pārsūtīšana no RAM uz datu glabāšanas ziņā efektīvu Apache Parquet formātu , izmantojot dažas Python rindas.

Nākamais piemērs: simulēsim ETL procesu ar datiem, kas glabājas parketa formātā, lai filtrētu projektus pēc noteiktas viena atribūta "cost\_million\_usd" vērtības ( 8.1-4. attēls).

❓ Kamēr tērzēšana turpinās, nosūtiet teksta pieprasījumu uz LLM:

Uzrakstiet kodu, kurā vēlamies filtrēt datus tabulā un saglabāt tikai tos projektus (tabulas rindas) no Apache Parquet datiem, kuru izmaksas (parametrs cost\_million\_usd) pārsniedz 150 miljonus ASV dolāru. ↴

💡 LLM atbilde :

```

1 # Loading construction project data
2 df_loaded_projects = pd.read_parquet(parquet_path_projects)
3
4 # Filtering out projects with a value greater than $150 million
5 df_filtered_projects = df_loaded_projects[df_loaded_projects['cost_million_usd'] > 150]
6
7 # Saving the filtered data to a new Parquet file
8 filtered_parquet_path_projects = '/tmp/construction_data/filtered_projects.parquet'
9 df_filtered_projects.to_parquet(filtered_parquet_path_projects)
10
11 print("Filtered construction projects data saved to Lakehouse.")

```

Run in IDE

**Attēls 8.1-5 ETL process , strādājot ar datiem Apache Parquet formātā, izskatās tāpat kā ar citiem strukturētiem formātiem.**

Izmantojot parketa formātu (salīdzinājumā ar XLSX, CSV u. c.), ievērojami samazina saglabātās informācijas apjomu un paātrina meklēšanas operācijas. Tāpēc tas ir lieliski piemērots gan datu glabāšanai, gan analīzei. Parquet integrējas ar dažādām apstrādes sistēmām, nodrošinot efektīvu piekļuvi hibrīdās arhitektūrās.

Tomēr efektīvs glabāšanas formāts ir tikai viens no pilnvērtīgas datu lietošanas pieredzes elementiem. Lai izveidotu ilgtspējīgu un mērogojamu vidi, ir nepieciešama labi izstrādāta datu pārvaldības arhitektūra. DWH (datu noliktavas) klasses sistēmas pilda šo funkciju. Tās nodrošina datu apkopošanu no heterogēniem avotiem, uzņēmējdarbības procesu pārredzamību un iespēju veikt sarežģītu analīzi, izmantojot BI rīkus un mašīnmācīšanās algoritmus.

## DWH: Datu noliktava datu noliktavas

Tāpat kā parketa formāts ir optimizēts efektīvai liela apjoma informācijas glabāšanai, datu noliktava ir optimizēta datu integrēšanai un strukturēšanai, lai atbalstītu analītiku, prognozēšanu un vadības lēmumu pieņemšanu.

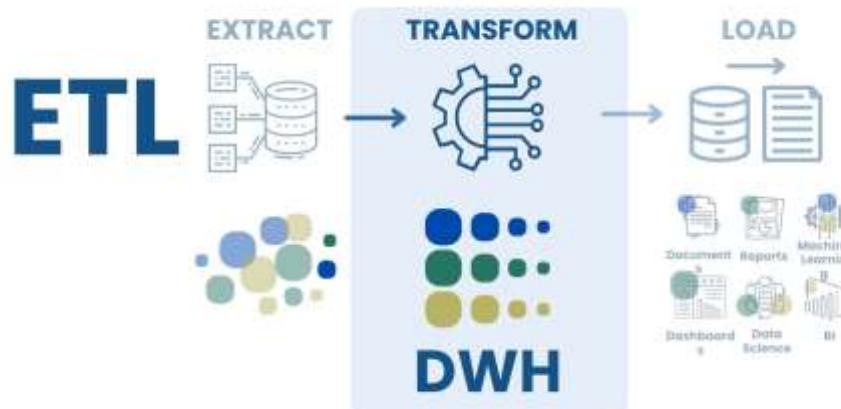
Mūsdienu uzņēmumos dati nāk no daudziem atšķirīgiem avotiem: ERP, CAFM, CPM, CRM sistēmām, grāmatvedības un noliktavu pārvaldības, ēku digitālajiem CAD modeļiem, IoT sensoriem un citiem risinājumiem. Lai iegūtu visaptverošu priekšstatu, nepietiek tikai apkopot datus - tie ir jāorganizē, jāstandartizē un jācentralizē vienotā krātuvē. DWH ir tieši tas, ko DWH dara - centralizēta glabāšanas sistēma, kas jauj apkopot informāciju no dažādiem avotiem, strukturēt to un padarīt pieejamu analīzei un stratēģiskai pārvaldībai.

**DWH (Data Warehouse)** ir centralizēta datu noliktavas sistēma, kas apkopo informāciju no dažādiem avotiem, strukturē to un dara pieejamu analītikai un pārskatu sagatavošanai.

Daudzos uzņēmumos dati ir izkaisīti pa dažādām sistēmām, par ko mēs runājām grāmatas pirmajās daļās ( 1.2-4. attēls). DWH integrē šos avotus, nodrošinot pilnīgu informācijas pārredzamību un uzticamību. DWH datu noliktava ir specializēta datu bāze (liela datu bāze), kas apkopo, apstrādā un glabā datus no vairākiem avotiem. DWH galvenās īpašības ir šādas:

- **Izmantojot ETL -procesus** (Extract, Transform, Load) - datu iegūšana no avotiem, to attīrīšana, pārveidošana, ielādēšana repositorijā un šo procesu automatizēšana, kas tika apskatīti grāmatas septītajā daļā.
- **Datu granularitāte** - datus DWH var glabāt gan apkopotā veidā (kopsavilkuma ziņojumi), gan granulārā veidā (neapstrādāti dati). Sākot ar 2024. gadu, tieši CAD- pārdevēji ir sākuši runāt par granulāriem datiem [125], kas, iespējams, norāda, ka nozare gatavojas pārejai uz specializētu mākoņglabāšanu, lai apstrādātu digitālos ēku modeļu datus.
- **Atbalsta analītika un prognozēšana** - datu noliktavas nodrošina pamatu BI rīkiem, lielo datu analīzei un mašīnmācīšanai.

DWH kalpo par pamatu biznesa inteliģencei, ņaujot analizēt galvenos darbības rādītājus, prognozēt pārdošanas, iepirkumu un izmaksu apjomu, kā arī sniegt automatizētus pārskatus un datu vizualizāciju ( 8.1-6. attēls).



Attēls 8.1-6 ETL -procesā DWH var darboties kā centrālā krātuve, kurā no dažādām sistēmām iegūtie dati tiek transformēti un izkrauti.

DWH ir galvenā loma informācijas integrēšanā, attīrīšanā un strukturēšanā, nodrošinot stabilu pamatu biznesa izlūkošanas un lēnumu pieņemšanas procesiem. Tomēr mūsdienā apstākļos, kad datu apjomi strauji pieaug un datu avoti kļūst daudzveidīgāki, tradicionālā DWH pieeja informācijas glabāšanai bieži vien prasa paplašināšanu ELT un datu ezera veidā

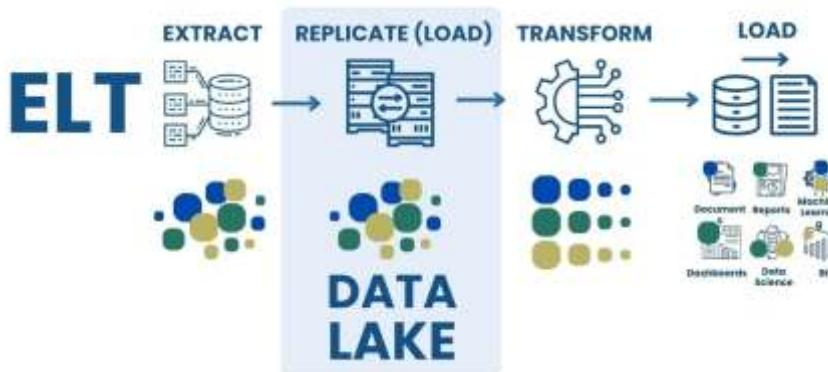
## Datu ezers - no ETL uz ELT: no tradicionālās tīrišanas uz elastīgu apstrādi

Klasiskās DWH - datu noliktavas, kas paredzētas strukturētu datu glabāšanai analītiskiem pieprasījumiem optimizētā formātā, ir saskārušās ar ierobežojumiem, apstrādājot nestukturētus datus un mērogojamību. Lai risinātu šīs problēmas, ir izveidoti datu ezeri), kas piedāvā elastīgu liela apjoma heterogēnu datu uzglabāšanu.

**Datu ezers** piedāvā alternatīvu DWH pieeju, kas ļauj strādāt ar nestukturētiem, daļēji strukturētiem un neapstrādātiem datiem bez iepriekšējas stingras shēmas. Šī glabāšanas metode bieži ir svarīga reāllaika datu apstrādei, mašīnmācīšanās un progresīvai analīzei. Atšķirībā no DWH, kas strukturē un apkopo datus pirms ielādes, datu ezers ļauj informāciju uzglabāt neapstrādātā veidā, tādējādi nodrošinot elastību un mērogojamību.

Tieši neapmierinātība ar tradicionālajām datu noliktavām (RDBMS, DWH) un interese par "lielajiem datiem" noveda pie datu ezeru rašanās, kur sarežģītas ETL vietā dati tagad tiek vienkārši augšupielādēti brīvi strukturētā repositorijā, bet apstrāde notiek analīzes posmā:

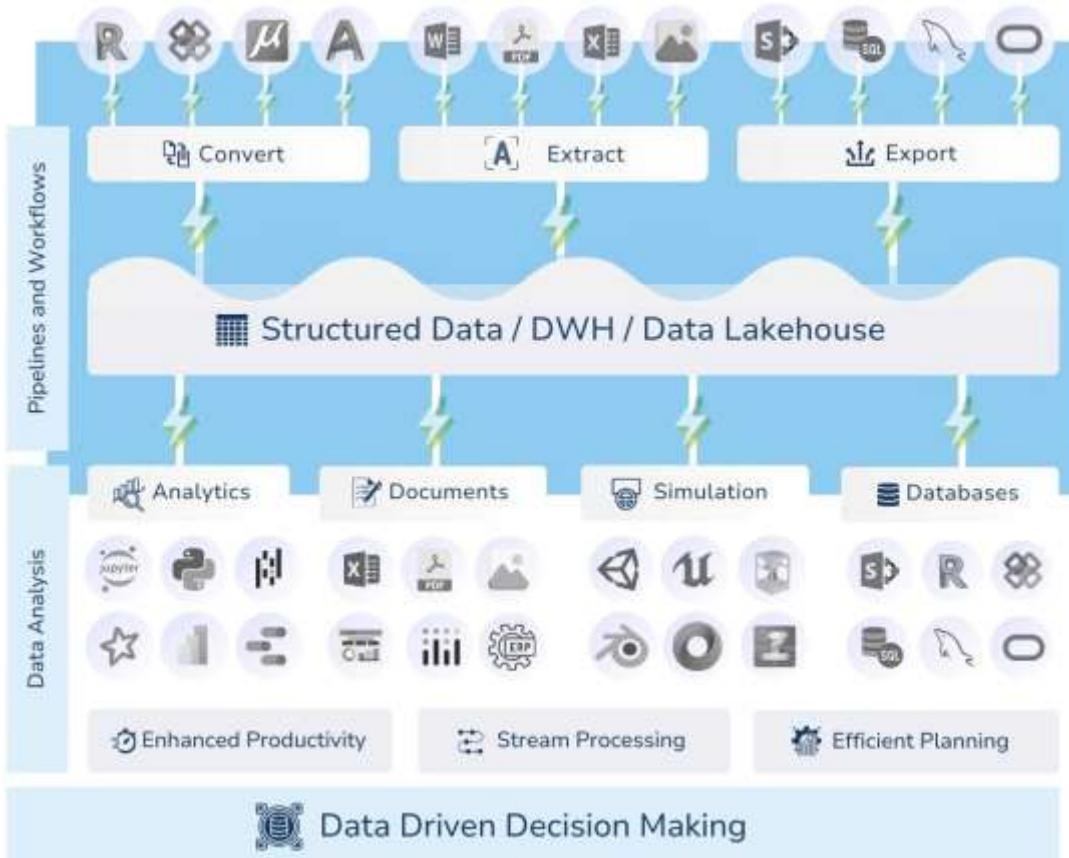
- Tradicionālajās datu noliktavās dati parasti tiek iepriekš apstrādāti, pārveidoti un attīrti (ETL - Extract, Transform, Load), pirms tie tiek ielādēti noliktavā (8.1-6. attēls). Tas nozīmē, ka dati tiek strukturēti un optimizēti konkrētiem turpmākiem analītiskiem un ziņošanas uzdevumiem. Uzsvars tiek likts uz augstas vaicājumu veikspējas un datu integritātes saglabāšanu. Tomēr šī pieeja var būt dārga un mazāk elastīga attiecībā uz jaunu datu tipu integrēšanu un strauji mainīgām datu shēmām.
- Datu ezeri, no otras pusēs, ir paredzēti liela apjoma neapstrādātu datu glabāšanai to sākotnējā formātā ( 8.1-7. attēls). ETL (Extract, Transform, Load), process tiek aizstāts ar ELT (Extract, Load, Transform), kur dati vispirms tiek ielādēti noliktavā "tādi, kādi ir", un tikai pēc tam tos var pārveidot un analizēt pēc vajadzības. Tas nodrošina lielāku elastību un iespēju uzglabāt heterogēnus datus, tostarp nestukturētus datus, piemēram, tekstu, attēlus un žurnālus.



Attēls 8.1-7 Atšķirībā no ETL, datu ezerā izmanto ELT, kurā informācija vispirms tiek augšupielādēta "neapstrādātā" veidā un augšupielādes posmā tiek veikta transformācija.

Tradicionālajās datu noliktavās galvenā uzmanība tiek pievērsta datu pirmapstrādei, lai nodrošinātu augstu meklēšanas veikspēju, savukārt datu ezeru prioritāte ir elastīgums: tie glabā neapstrādātus

datus un pārveido tos pēc vajadzības (8.1.1-8. attēls).



Attēls 8.1-8 Mūsdienu glabāšanas koncepcijas mērķis ir uzglabāt un apstrādāt visu veidu datus lēmumu pieņemšanas vajadzībām.

Tomēr, neraugoties uz visām priekšrocībām, datu ezeri nav bez trūkumiem. Stingras struktūras trūkums un informācijas pārvaldības sarežģītība var radīt haosu, kurā dati dublējas, ir pretrunīgi vai kļūst nebūtiski. Turklāt datu meklēšana un analīze šādā krātuvē prasa ievērojamas pūles, īpaši, ja runa ir par neviendabīgu informāciju. Lai pārvarētu šos ierobežojumus un apvienotu labākās tradicionālo datu noliktavu un datu ezeru īpašības, tika izstrādāta datu ezeru arhitektūra.

### Datu ezernīcas arhitektūra: noliktavu un datu ezeru sinergija

Lai apvienotu labākās DWH (strukturēta, pārvaldāma, augstas veikspējas analītika) un datu ezera (mērogojamība, heterogēnu datu apstrāde) īpašības, tika izstrādāta datu ezera pieeja. Šī arhitektūra apvieno datu ezeru elastīgumu ar jaudīgiem apstrādes un pārvaldības rīkiem, kas raksturīgi tradicionālajām noliktavām, panākot līdzsvaru starp glabāšanu, analītiku un mašīnmācīšanos. Data Lakehouse ir datu ezeru un datu noliktavu sintēze, kas apvieno datu ezeru elastību un mērogojamību ar datu noliktavu pārvaldības un vaicājumu optimizācijas iespējām.

**Datu ezeri** ir arhitektūras pieeja, kuras mērķis ir apvienot datu ezeru elastību un mērogojamību ar datu noliktavu pārvaldāmību un vaicājumu veikspēju ( 8.1-9. attēls).

Datu ezernīcas galvenās iezīmes:

- **Atvērts datu glabāšanas formāts:** atklātu datu glabāšanas formātu, piemēram, Apache Parquet, izmantošana nodrošina efektivitāti un optimizētus pieprasījumus.
- Tikai lasīšanai paredzēta **shēma:** atšķirībā no DWH tradicionālās shēmas, kas paredzēta tikai rakstīšanai, Lakehouse atbalsta tikai lasīšanai paredzētu shēmu, kas ļauj elastīgāk pārvaldīt datu struktūru.
- **Elastīgs un mērogojams:** atbalsta strukturētu un nestukturētu datu glabāšanu un analīzi, nodrošinot augstu meklēšanas veikspēju, izmantojot glabāšanas līmeņa optimizāciju.

Data Lakehouse piedāvā kompromisa risinājumu, kas apvieno abu pieju priekšrocības, padarot to ideāli piemērotu mūsdienīgām analītiskajām slodzēm, kurām nepieciešama elastīga datu apstrāde.



Attēls 8.1-9 Data Lakehouse ir nākamās paudzes glabāšanas sistēmas, kas izstrādātas, lai atbilstu sarežģītām un pastāvīgi mainīgām prasībām.

Mūsdienu datu noliktavu ideja šķiet vienkārša: ja visi dati ir vienuviet, tos ir vieglāk analizēt. Tomēr praksē viss nav tik gludi. Iedomājieties, ka uzņēmums nolej pilnībā atteikties no ierastajām grāmatvedības un vadības sistēmām (ERP, PMIS, CAFM vai citām), aizstājot tās ar vienu milzīgu datu ezeru, kam var piekļūt visi. Kas notiks? Visticamāk, iestāsies haoss: dati dublēsies, būs pretrunīgi, kritiski svarīga informācija tiks pazaudēta vai bojāta. Pat ja datu ezeru izmantos tikai analītikai, bez pienācīgas pārvaldības tas būs nopietni apdraudēts:

- Datus ir grūti izprast: parastajās sistēmās datiem ir skaidra struktūra, bet ezerā tie ir tikai milzīgs failu un tabulu kopums. Lai kaut ko atrastu, ir jānoskaidro, par ko ir atbildīga katra rinda un sleja.

- Dati var būt neprecīzi: ja vienā vietā tiek glabātas vairākas vienas un tās pašas informācijas versijas, ir grūti noteikt, kura versija ir aktuāla. Rezultātā lēmumi tiek pieņemti, pamatojoties uz novecojušiem vai klūdainiem datiem.
- Ir grūti sagatavot datus darbam: dati ne tikai jāuzglabā, bet arī jāsniedz ērtā formā - atskaišu, grafiku, tabulu veidā. Tradicionālajās sistēmās tas tiek darīts automātiski, bet datu ezeros tas prasa papildu apstrādi.

Rezultātā katrai datu glabāšanas koncepcijai ir savas īpatnības, apstrādes pieejas un biznesa lietojumi. Tradicionālās datubāzes koncentrējas uz transakciju operācijām, datu noliktavas (DWH) nodrošina struktūru analītikai, datu ezeri (Data Lake) glabā informāciju neapstrādātā veidā, bet hibrīdās noliktavas (Data Lakehouse) apvieno DWH un Data Lake priekšrocības (8.1-10. attēls).

	Traditional Approach	Data Warehouse	Data Lake	Data Lakehouse
Data Types	Relational Databases	Structured, ready for analytics	Raw, semi-structured, or unstructured	Mix of structured and unstructured
Use Cases	Transactional Systems	Reporting, dashboards, BI	Big data storage, AI, advanced analytics	Hybrid analytics, AI, real-time data
Processing	OLTP – real-time transactions	ETL – clean and structure before analysis	ELT – store raw data, transform later	ELT with optimized storage and real-time processing
Storage	On-premise servers	Centralized, SQL-based	Decentralized, flexible formats	Combines advantages of DWH and DL
Common Tools	MySQL, PostgreSQL	Snowflake, Redshift, BigQuery	Hadoop, AWS S3, Azure Data Lake	Databricks, Snowflake, Google BigLake

Attēlā 8.1-10. DWH, datu ezers un datu ezers: galvenās atšķirības datu tipos, izmantošanas scenārijos, apstrādes metodēs un glabāšanas pieejās.

Glabāšanas arhitektūras izvēle ir sarežģīts process, kas atkarīgs no biznesa vajadzībām, informācijas apjoma un analītiskajām prasībām. Katram risinājumam ir savi plusi un mīnusi: DWH nodrošina struktūru, Data Lake nodrošina elastību, bet Lakehouse nodrošina līdzsvaru starp abiem risinājumiem. Organizācijas reti aprobežojas ar vienu datu arhitektūru.

Neatkarīgi no izvēlētās arhitektūras automatizētās datu pārvaldības sistēmas ir ievērojami pārākas par manuālajām metodēm. Tās līdz minimumam samazina cilvēka klūdu skaitu, paātrina informācijas apstrādi un nodrošina datu pārredzamību un izsekojamību visos uzņēmējdarbības procesu posmos.

Lai gan centralizētās datu noliktavas jau ir kļuvušas par nozares standartu daudzās ekonomikas jomās, situācija būvniecības nozarē joprojām ir sadrumstalota. Dati šeit ir izkliedēti dažādās platformās (CDE, PMIS, ERP u. c.), kas apgrūtina vienota priekštata veidošanu par notiekošo un prasa arhitektūras, kas spēj apvienot šos avotus vienotā, analītiski izmantojamā digitālā vidē.

## CDE, PMIS, ERP vai DWH un datu ezers.

Daži būvniecības un inženiertehniskie uzņēmumi jau izmanto kopējās datu vides (CDE) koncepciju saskaņā ar ISO 19650. Būtībā CDE pilda tādas pašas funkcijas kā datu noliktava (DWH) citās nozarēs: centralizē informāciju, nodrošina versiju kontroli, nodrošina piekļuvi apstiprinātai informācijai.

**Kopējā datu vide (CDE)** ir centralizēta digitāla telpa, ko izmanto, lai pārvaldītu, uzglabātu, kopīgotu un sadarbotos ar projekta informāciju visos objekta dzīves cikla posmos. CDE bieži tiek ieviesta, izmantojot mākoņtehnoloģijas, un integrēta ar CAD (BIM) sistēmām.

Finanšu, mazumtirdzniecības, loģistikas un rūpniecības nozarēs jau vairākus gadus desmitus tiek izmantotas centralizētas datu pārvaldības sistēmas, kas apvieno informāciju no dažādiem avotiem, kontrolē tās nozīmīgumu un nodrošina analīzi. CDE šos principus turpina attīstīt, pielāgojot tos ēku projektēšanas un dzīves cikla pārvaldības problēmām.

Tāpat kā DWH, CDE strukturē datus, fiksē izmaiņas un nodrošina vienotu piekļuves punktu pārbaudītai informācijai. Līdz ar pāreju uz mākoņdatošanu un integrāciju ar analītiskajiem rīkiem atšķirības starp abiem klūst arvien mazāk pamānāmas. Pievienojot CDE granulārus datus, par kuru koncepciju CAD - pārdevēji ir diskutējuši jau kopš 2023. gada[93, 125], var saskatīt vēl vairāk paralēlu ar klasisko DWH.

Iepriekš nodaļā "Būvniecības ERP un PMIS sistēmas" mēs jau apskatījām PMIS (projektu vadības informācijas sistēmu) un ERP (uzņēmumu resursu plānošanas sistēmu). Būvniecības projektos CDE un PMIS darbojas kopā: CDE kalpo kā datu, tostarp rasējumu, modeļu un projekta dokumentācijas, krātuve, savukārt PMIS pārvalda tādus procesus kā termiņu, uzdevumu, resursu un budžetu kontrole.

ERP, kas ir atbildīga par uzņēmējdarbības pārvaldību kopumā (finances, iepirkumi, personāls, ražošana), var integrēties ar PMIS, nodrošinot izmaksu un budžeta kontroli uzņēmuma līmenī. Analītikai un ziņošanai DWH var izmantot, lai apkopotu, strukturētu un apkopotu datus no CDE, PMIS un ERP, lai novērtētu finanšu KPI (ROI) un noteiku modeļus. Savukārt datu ezers (DL) var papildināt DWH, uzglabājot neapstrādātus un nestukturētus datus (piemēram, žurnālus, sensoru datus, attēlus). Šos datus var apstrādāt un ielādēt DWH tālākai analīzei.

Tādējādi CDE un PMIS pievēršas projektu vadībai, ERP - biznesa procesiem, bet DWH un Data Lake - analītikai un datu operācijām.

Saīdzinot CDE, PMIS un ERP sistēmas ar DWH un datu ezeru, var novērot būtiskas atšķirības attiecībā uz piegādātāja neatkarību, izmaksām, integrācijas elastību, datu neatkarību, pielāgošanās ātrumu pārmaiņām un analītiskajām iespējām (8.1-11. attēls). Tradicionālās sistēmas, piemēram, CDE, PMIS un ERP, bieži vien ir piesaistītas konkrētiem piegādātāju risinājumiem un standartiem, kas padara tās mazāk elastīgas un palielina to izmaksas licenču un atbalsta dēļ. Turklāt dati šādās sistēmās bieži vien ir iekapsulēti patentētos, slēgtos formātos, kas ierobežo to izmantošanu un analīzi.

		CDE, PMIS, ERP	DWH, Data Lake
	Vendor Dependency	High (tied to specific solutions and standards of vendors)	Low (flexibility in tool and platform choice)
	Integration Flexibility	Limited (integration depends on vendor solutions)	High (easily integrates with various data sources)
	Cost	High (licensing and support costs)	Relatively lower (use of open technologies and platforms)
	Data Independence	Low (data often locked in proprietary formats)	High (data stored in open and accessible formats)
	Adaptability to Changes	Slow (changes require vendor approval and integration)	Fast (adaptation and data structure modification without intermediaries)
	Analytical Capabilities	Limited (dependent on vendor-provided solutions)	Extensive (support for a wide range of analytical tools)

**Attēlā: 8.1-11 DWH un datu ezers piedāvā lielāku elastību un datu neatkarību nekā tādas sistēmas kā CDE, PMIS un ERP.**

Turpretī DWH un Data Lake nodrošina lielāku elastību integrācijā ar dažādiem datu avotiem, un to atvērto tehnoloģiju un platformu izmantošana palīdz samazināt kopējās īpašumtiesību izmaksas. Turklāt DWH un Data Lake atbalsta plašu analītisko rīku klāstu, kas uzlabo analīzes un pārvaldības iespējas.

Attīstoties CAD formātu reversās inženierijas rīkiem un pieķuvei CAD lietojumprogrammu datubāzēm, arvien aktuālāks kļūst jautājums: cik pamatoti ir turpināt izmantot slēgtas, izolētas platformas, ja projektēšanas datiem jābūt pieejamiem plašam speciālistu lokam, kas strādā desmitiem darbuzņēmēju un projektēšanas organizāciju?

Šāda atkarība no konkrēta piegādātāja tehnoloģijas var ievērojami ierobežot datu pārvaldības elastību, palēnināt reakciju uz projekta izmaiņām un kavēt efektīvu sadarbību starp dalībniekiem.

Tradicionālās pieejas datu pārvaldībai - tostarp DWH, datu ezers, CDE un PMIS - galvenokārt ir vērstas uz informācijas glabāšanu, strukturēšanu un apstrādi. Tomēr, attīstoties mākslīgajam intelektam un mašīnmācīšanai, pieaug vajadzība pēc jauniem datu organizēšanas veidiem, kas ne tikai apkopo, bet arī atklāj sarežģītas sakārības, atrod slēptos modeļus un nodrošina tūlītēju piekļuvi vissvarīgākajai informācijai.

Īpaša loma šajā virzienā sāk būt vektoru datubāzēm - jauna veida krātuvēm, kas optimizētas lieldimensiju iestrādnēm.



## NODĀĻA 8.2.

### DATU NOLIKTAVAS PĀRVALDĪBA UN HAOSA NOVĒRŠANA

#### Vektoru datubāzes un robežlaukums

Vektoru datubāzes ir jauna krātuvju klase, kas ne tikai glabā datus, bet arī ļauj veikt meklēšanu pēc nozīmes, salīdzināt objektus pēc semantiskā tuvuma un veidot inteliģentas sistēmas - no ieteikumiem līdz automātiskai analīzei un konteksta ģenerēšanai. Atšķirībā no tradicionālajām datubāzēm, kas koncentrējas uz precīzām sakritībām, vektoru datubāzes atrod līdzīgus objektus, pamatojoties uz atribūtiem, pat ja nav precīzas sakritības

**Vektoru** datubāze ir specializēts datubāzes veids, kurā dati tiek glabāti kā daudzdimensiju vektori, no kuriem katrs pārstāv noteiktas īpašības vai īpašības. Šiem vektoriem var būt dažāds dimensiju skaits atkarībā no datu sarežģītības (vienā gadījumā tās var būt dažas dimensijas, bet citā — tūkstoši).

Galvenā vektoru datubāzu priekšrocība ir meklēšana pēc semantiskās atbilstības, nevis pēc precīzas vērtību atbilstības. SQL un Pandas -ieprasījumu ar "equals" vai "contains" filtriem vietā tiek izmantota tuvāko kaimiņu (k-NN) meklēšana (par k-NN plašāk runāsim nākamajā grāmatas daļā) pazīmju telpā.

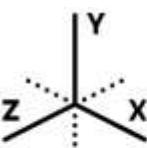
Līdz ar LLM (Large Language Models) un ģeneratīvo modeļu attīstību mijiedarbība ar datubāzēm sāk mainīties. Tagad ir iespējams veikt datu pieprasījumus dabiskajā valodā, semantiski meklēt dokumentus, automātiski iegūt atslēgas terminus un veidot kontekstuālas attiecības starp objektiem - un tas viss bez nepieciešamības pārzināt SQL vai pārzināt tabulu struktūru. Par to sīkāk tika runāts sadaļā "LLM un to loma datu apstrādē un uzņēmējdarbības procesos".

Tomēr ir svarīgi apzināties, ka LLM automātiski nestrukturē un nesakārto informāciju. Modelis vienkārši izplūst cauri datiem un atrod visatbilstošāko datu daļu, pamatojoties uz vaicājuma kontekstu. Ja dati nav iepriekš attīrti vai pārveidoti, padziļinātā meklēšana būs līdzīga mēģinājumam atrast atbildi digitālajā "atkritumos" - tas var izdoties, bet rezultātu kvalitāte būs zemāka. Ideālā gadījumā, ja datus var strukturēt (piemēram, tulkot dokumentus Markdown formātā) un ielādēt vektoru datubāzē. Tas ievērojami palielina rezultātu precizitāti un atbilstību.

Sākotnēji vektoru datubāzes tika izmantotas mašīnmācībā, taču šodien tās arvien vairāk izmanto arī ārpus tās - meklētājprogrammās, saturu personalizācijā un inteliģentajā analīzē.

Viens no acīmredzamākajiem vektoru piejas piemēriem konstruēšanā ir ierobežojošais lodziņš (ierobežojošais paralēlālemeħānisms). Tā ir ģeometriskā konstrukcija, kas apraksta objekta robežas trīsdimensiju telpā. Bounding Box definē minimālās un maksimālās X, Y un Z koordinātas, veidojot "kastīti" ap objektu. Šī metode ļauj novērtēt elementa izmēru un izvietojumu, neanalizējot visu ģeometriju.

Katru robežstūri var attēlot kā vektoru daudzdimensiju telpā: piemēram, [x, y, z, platoms, augstums, dziļums] - jau 6 dimensijas ( 8.2-1. attēls).



## Bounding Box

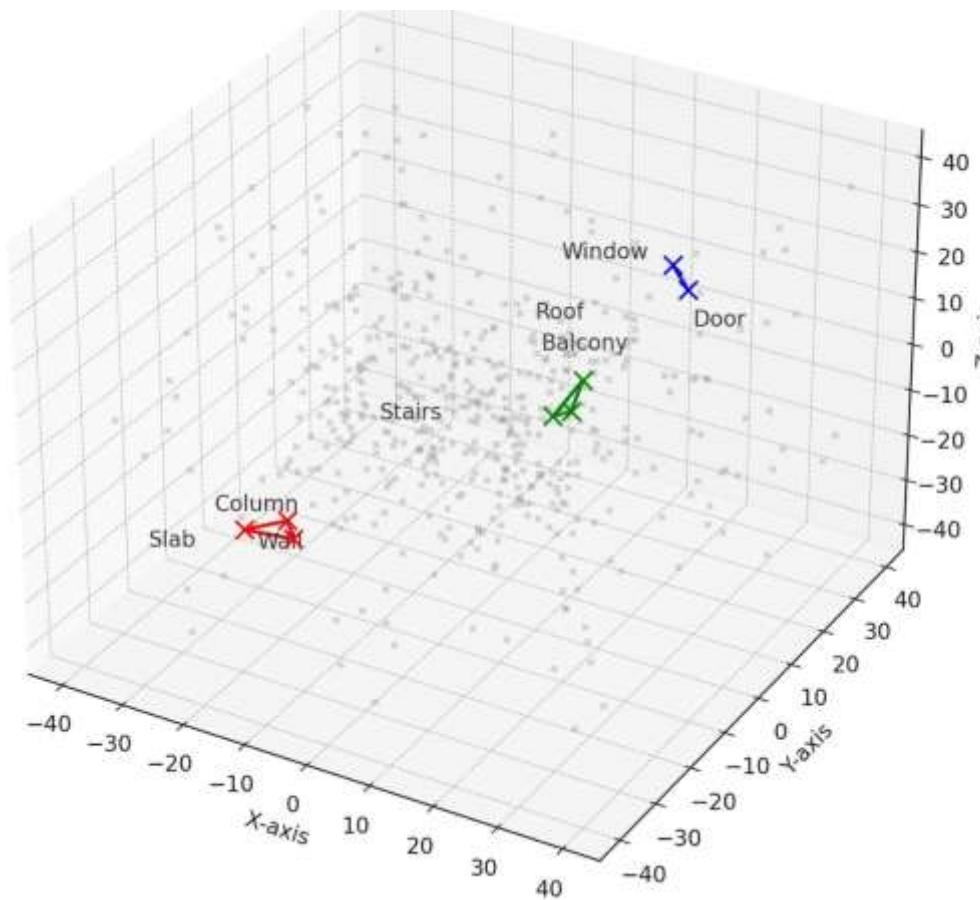
	minX	maxX	minY	maxY	minZ	maxZ	Width	Height	Depth
Column	-15	-5	-25	-15	0	10	10	10	20
Stairs	-5	5	-15	-5	0	10	10	10	10
Door	5	15	5	15	0	10	10	10	10
Window	25	35	-35	-25	10	30	10	20	20
Balcony	15	25	-5	5	20	40	10	20	20

Attēls 8.2-1 Bounding Box -elementu koordinātu informācija un to atrašanās vieta projekta modelī ir analoga vektoru datubāzei.

Šāda datu attēlojums atvieglo daudzus uzdevumus, tostarp objektu krustpunktu pārbaudi, ēkas elementu telpiskā izvietojuma plānošanu un automatizētu aprēķinu veikšanu. Bounding Box var kalpot kā tilts starp sarežģītiem 3D modeļiem un tradicionālajām vektoru datubāzēm, lāujot efektīvi izmantot abu pieeju priekšrocības arhitektūras un inženiertehniskajā modelēšanā.

Bounding Box ir "ģeometrijas vektorizācija", un embedding (veids, kā pārveidot kaut ko abstraktu) ir "nozīmes vektorizācija". Abas pieejas lāuj pāriet no manuālas meklēšanas uz intelīgentu meklēšanu, neatkarīgi no tā, vai runa ir par 3D -objektiem projekta modelī vai jēdzieniem tekstā.

Objektu meklēšana projektā (piemēram, "atrast visus logus, kuru platoms > 1,5 m") ir līdzīga tuvāko kaimiņu meklēšanai (k-NN) vektoru datubāzē, kur kritēriji nosaka "zonu" pazīmju telpā. (vairāk par k-NN tuvāko kaimiņu meklēšanu mēs runāsim nākamajā daļā par mašīnmācīšanos) ( 8.2-2. attēls). Ja norobežojošās zonas atribūtiem pievienojam papildu parametrus (materiāls, svars, ražošanas laiks), tabula pārvēršas par daudzdimensiju vektoru, kurā katrs atribūts ir jauna dimensija. Tas ir tuvāk mūsdienu vektoru bāzēm, kur dimensijas tiek skaitītas simtos vai tūkstošos (piemēram, iestrādājot no neironu tīkliem).



Attēls 8.2-2 Objektu meklēšana projektā, izmantojot vektoru datubāzes.

Pieeja, kas izmantota Bounding Box, ir piemērojama ne tikai ģeometriskiem objektiem, bet arī teksta un valodas analīzei. Datu vektoru attēlojumus jau aktīvi izmanto dabiskās valodas apstrādē (NLP). Tāpat kā objektus būvprojektā var sagrupēt pēc to telpiskā tuvuma (8.2-2. attēls), arī vārdus tekstā var analizēt pēc to semantiskā un kontekstuālā tuvuma.

Piemēram, vārdi "arhitekts", "būvniecība", "dizains" atrodas blakus viens otram vektoru telpā, jo tiem ir līdzīga nozīme. LLM šis mehānisms ļauj automātiski, bez nepieciešamības veikt manuālu kategorizāciju:

- Identificēt teksta tēmu
- Veikt semantisko meklēšanu dokumentu saturā.
- Automātisku anotāciju un teksta kopsavilkumu ģenerēšana
- Atrodiet sinonīmus un saistītus terminus

Vektoru datubāzes ļauj analizēt tekstu un atrast tajā saistītus terminus tādā pašā veidā, kā Bounding Box palīdz analizēt telpiskus objektus 3D-modeļos. Bounding Box piemērs ar projekta elementiem palīdz saprast, ka vektoru attēlojums nav tikai "mākslīgs" jēdziens no ML, bet gan dabisks datu strukturēšanas veids lietišķo problēmu risināšanai, neatkarīgi no tā, vai tā ir kolonnu meklēšana CAD

projektā vai semantiski tuvu attēlu meklēšana datubāzē.

Speciālistiem, kas strādā ar datubāzēm, jāpievērš uzmanība vektoru veikaliem. To izplatība liecina par jaunu posmu datubāzu attīstībā, kurā klasiskās relāciju sistēmas un uz mākslīgo intelekta orientētās tehnoloģijas sāk savstarpēji mijiedarboties, veidojot nākotnes hibrīdrisinājumus.

Lietotāji, kas izstrādā sarežģītas un liela mēroga mākslīgā intelekta lietojumprogrammas, izmantos specializētas datubāzes vektoru meklēšanai. Tajā pašā laikā tie, kam vajadzīgas tikai atsevišķas AI funkcijas integrēšanai esošajās lietojumprogrammās, visticamāk, izvēlēsies jau izmantotajās datubāzēs (PostgreSQL, Redis) iebūvētas vektoru meklēšanas iespējas.

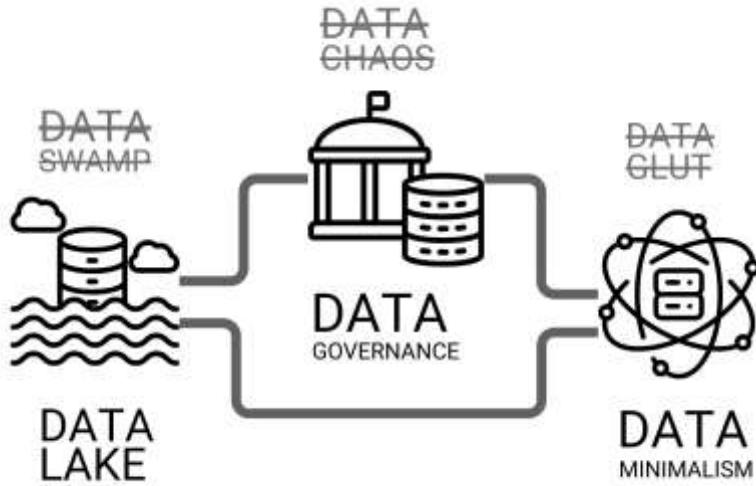
Lai gan tādas sistēmas kā DWH, datu ezers, CDE, PMIS, vektoru datubāzes un citas piedāvā dažādas pieejas datu glabāšanai un pārvaldībai, to efektivitāti nosaka ne tikai to arhitektūra, bet arī tas, cik labi tiek organizēti un pārvaldīti paši dati. Pat tad, ja tiek izmantoti mūsdienīgi risinājumi - vai tās būtu vektoru datubāzes, klasiskās relāciju DBVS vai Data Lake tipa noliktavas - skaidru datu pārvaldības, strukturēšanas un atjaunināšanas noteikumu trūkums var radīt tās pašas grūtības, ar kurām saskaras lietotāji, kas strādā ar atšķirīgām datnēm un dažāda formāta datiem.

Bez datu pārvaldības) pat visefektīvākie risinājumi var kļūt haotiski un nestukturēti, pārvēršot datu ezerus par datu purviem). Lai no tā izvairītos, uzņēmumiem ne tikai jāizvēlas pareizā glabāšanas arhitektūra, bet arī jāīsteno datu minimālisma), piekļuves pārvaldības un kvalitātes kontroles stratēģijas, lai datus pārvērstu par efektīvu lēmumu pieņemšanas rīku.

### **Datu pārvaldība), datu minimālisms) un datu purvs).**

Datu pārvaldības), datu minimālisma) un datu purva novēršanas) koncepciju izpratne un īstenošana ir galvenie priekšnosacījumi, lai veiksmīgi pārvaldītu datu noliktavas un nodrošinātu uzņēmējdarbības vērtību ( 8.2-3. attēls).

Saskaņā ar Gartner pētījumu (2017) 85% lielo datu projektu cieš neveiksni, un viens no galvenajiem iemesliem ir nepietiekama datu kvalitāte un datu pārvaldība [144].



**Attēls. 8.2-3 Daži no galvenajiem datu pārvaldības aspektiem ir datu pārvaldība un datu minimālisms.**

Datu pārvaldība (Data Governance) ir datu pārvaldības pamatkomponents, kas nodrošina, ka dati tiek izmantoti atbilstoši un efektīvi visos uzņēmējdarbības procesos. Tā ir ne tikai noteikumu un procedūru noteikšana, bet arī datu pieejamības, uzticamības un drošības nodrošināšana:

- datu definēšana un klasificēšana: skaidra vienību definēšana un klasificēšana ļauj organizācijām saprast, kādas vienības ir nepieciešamas uzņēmumā, un noteikt, kā tās jāizmanto.
- Piekļuves tiesības un pārvaldība: datu piekļuves un pārvaldības politikas un procedūru izstrāde nodrošina, ka konkrētiem datiem var piekļūt tikai pilnvarotī lietotāji.
- Datu aizsardzība pret ārējiem apdraudējumiem: datu aizsardzība pret ārējiem apdraudējumiem ir būtisks datu pārvaldības aspeks. Tas ietver ne tikai tehniskus pasākumus, bet arī darbinieku apmācību par informācijas drošības pamatiem.

Datu minimālisms (Data Minimalism) ir pieeja, kuras mērķis ir samazināt datus līdz vērtīgākajiem un nozīmīgākajiem atribūtiem un vienībām formējumā ( 8.2-4. attēls), tādējādi samazinot izmaksas un uzlabojot datu izmantošanu:

- Lēmumu pieņemšanas vienkāršošana: objektu un to atribūtu skaita samazināšana līdz svarīgākajiem objektiem vienkāršo lēmumu pieņemšanu, samazinot datu analīzei un apstrādei nepieciešamo laiku un resursus.
- Koncentrēšanās uz svarīgāko: atlasot būtiskākās vienības un atribūtus, varat koncentrēties uz informāciju, kas patiesām ir svarīga uzņēmumam, novēršot trokšņus un nevajadzīgus datus.
- Efektīva resursu piešķiršana: datu minimizēšana ļauj efektīvāk piešķirt resursus, samazinot datu glabāšanas un apstrādes izmaksas, uzlabojot datu kvalitāti un drošību.

Darba ar datiem loģikai jāsākas nevis ar to radīšanu kā tādu ( 8.2-4. attēls), bet gan ar izpratni par šo datu izmantošanas nākotnes scenārijiem vēl pirms datu ģenerēšanas procesa uzsākšanas. Šāda pieeja ļauj iepriekš definēt minimālās nepieciešamās prasības atribūtiem, to tipiem un robežvērtībām.

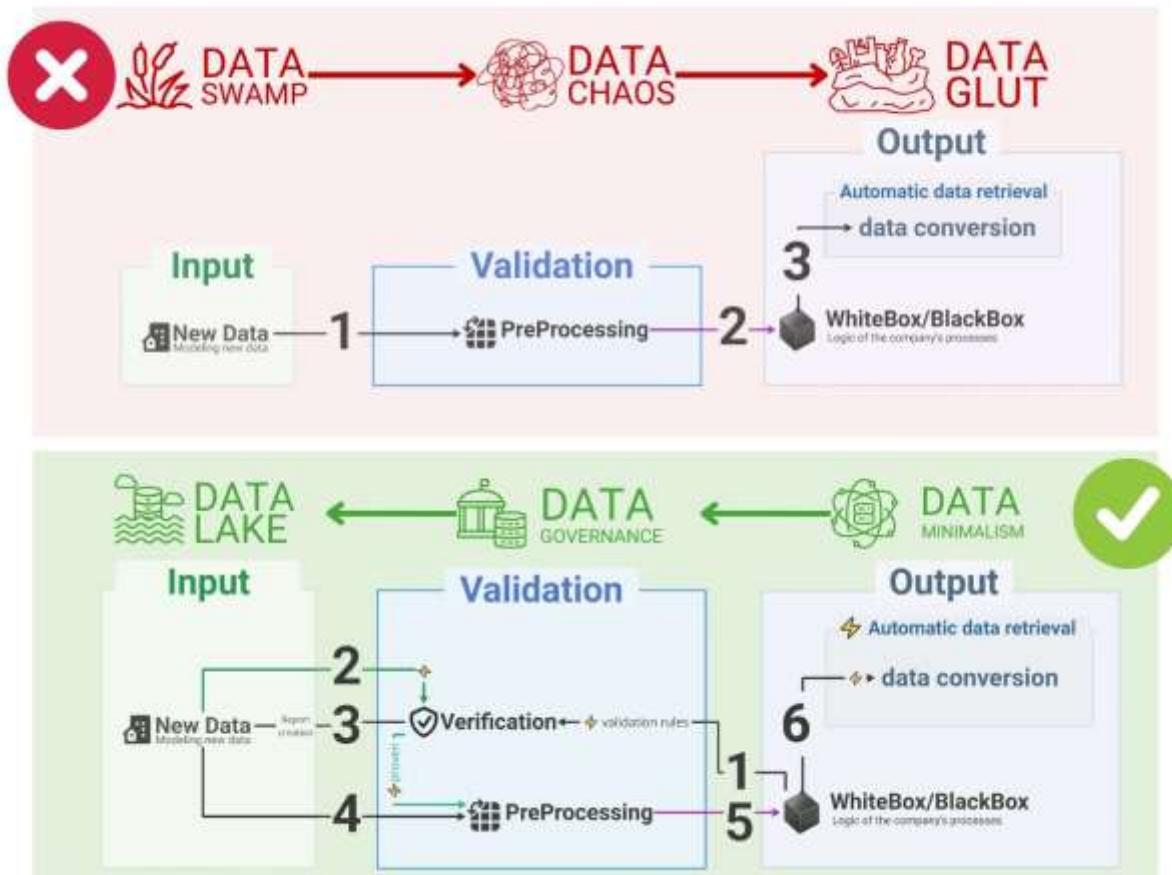
Šīs prasības veido pamatu pareizu un stabili vienību izveidei informācijas modelī. lepriekšēja izpratne par datu mērķiem un lietojumiem palīdz veidot analīzei piemērotu struktūru. Sīkāka informācija par datu modelēšanas pieejām konceptuālajā, loģiskajā un fiziskajā līmenī tika aplūkota nodaļā "Datu modelēšana: konceptuālais, loģisks un fizisks modelis".

Tradicionālajos būvniecības uzņēmumu biznesa procesos datu apstrāde biežāk atgādina datu izgāšanu purvā, kur dati vispirms tiek radīti un pēc tam speciālisti cenšas tos integrēt citās sistēmās un rīkos.

Datu purvs (Data Swamp) ir rezultāts nekontrolētai datu vākšanai un uzglabāšanai bez pienācīgas organizācijas, strukturēšanas un pārvaldības, kā rezultātā dati ir nestukturēti, grūti izmantojami un mazvērtīgi.

Kā nepieļaut, ka informācijas plūsma klūst par purvu:

- **Datu struktūras pārvaldība:** datu strukturēšanas un kategorizēšanas nodrošināšana palīdz novērst datu pārblīvēšanos, padarot tos sakārtotus un viegli pieejamus.
- **Datu izpratne un interpretācija:** skaidrs datu izcelsmes, modifikāciju un nozīmes apraksts nodrošina, ka dati tiek pareizi saprasti un interpretēti.
- **Datu kvalitātes uzturēšana:** regulāra datu uzturēšana un tīrišana palīdz uzturēt datu kvalitāti, atbilstību un vērtību analītikai un uzņēmējdarbības procesiem.



Attēls 8.2-4 Lai izvairītos no jucekļa datu krātuvē, datu izveides process jāsāk ar atribūtu prasību apkopošanu.

Integrējot datu pārvaldības un datu minimālisma principus datu pārvaldības procesos un aktīvi novēršot datu noliktavu pārvēršanos par datu purviem, organizācijas var maksimāli izmantot savu datu potenciālu.

Nākamais darba ar datiem attīstības posms pēc pārvaldības un minimālisma jautājumu atrisināšanas ir automātiskās apstrādes standartizācija, kvalitātes nodrošināšana un tādu metožu ieviešana, kas padara datus izmantojamus analīzei, pārveidošanai un lēmumu pieņemšanai. To dara DataOps un VectorOps metodoloģijas, kas klūst par svarīgiem rīkiem uzņēmumiem, kuri strādā ar lielajiem datiem un mašīnmācīšanos.

## DataOps un VectorOps: jaunie datu standarti

Datu pārvaldība ir atbildīga par datu kontroli un organizēšanu, savukārt DataOps palīdz nodrošināt datu precizitāti, konsekvenči un vienmērīgu plūsmu uzņēmumā. Tas ir īpaši svarīgi vairākiem uzņēmējdarbības gadījumiem būvniecībā, kur dati tiek generēti nepārtraukti un tos nepieciešams savlaicīgi apstrādāt. Piemēram, situācijās, kad būvniecības informācijas modeli, projekta prasības un

analītiskie pārskati jāsinhronizē starp dažādām sistēmām vienas darba dienas laikā, DataOps loma var būt ļoti svarīga. Tas ļauj izveidot stabilus un atkārtojamus datu apstrādes procesus, samazinot kavēšanās un informācijas aktualitātes zuduma risku.

Ar datu pārvaldību vien nepietiek - ir svarīgi, lai dati tiktu ne tikai glabāti, bet arī aktīvi izmantoti ikdienas darbībā. Tieši šeit ir jāievieš DataOps - metodoloģija, kas koncentrējas uz automatizāciju, integrāciju un nepārtrauktu datu plūsmu.

**DataOps** koncentrējas uz sadarbības uzlabošanu, integrāciju un datu plūsmu automatizāciju organizācijās. DataOps prakses pieņemšana veicina datu precizitāti, konsekvenci un pieejamību, kas ir ļoti svarīgi uz datiem orientētām lietojumprogrammām.

Galvenie rīki DataOps ekosistēmā ir Apache Airflow ( 7.4-4. attēls) darbplūsmu orķestrēšanai un Apache NiFi ( 7.4-5. attēls) datu plūsmu maršrutēšanai un transformēšanai. Kopā šīs tehnoloģijas nodrošina elastīgus, uzticamus un mērogojamus datu cauruļvadus, lai automātiski apstrādātu, kontrolētu un integrētu informāciju starp sistēmām (vairāk informācijas nodaļā "Automātiskais ETL - konveijers "). Īstenojot DataOps pieju būvniecības procesos, ir svarīgi ņemt vērā četrus pamataspektus:

1. **Cilvēki un rīki ir svarīgāki par datiem:** atsevišķas datu krātuves var uzskatīt par galveno problēmu, taču realitāte ir sarežģītāka. Papildus datu sadrumstalotībai liela nozīme ir arī komandu izolētībai un atšķirīgajiem rīkiem, ko tās izmanto. Būvniecībā ar datiem strādā dažādu jomu speciālisti: datu inženieri un analītiķi, BI un vizualizācijas komandas, kā arī projektu vadības un kvalitātes eksperti. Katram no viņiem ir atšķirīgi darba paņēmieni, tāpēc klūst svarīgi izveidot ekosistēmu, kurā dati brīvi plūst starp daļībniekiem, nodrošinot vienotu, konsekventu informācijas versiju.
2. **Automatizēt testēšanu un klūdu atklāšanu:** būvniecības datos vienmēr ir klūdas - neprecizitātes modeļos, aprēķinu klūdas vai novecojušas specifikācijas. Regulāra datu testēšana un atkārtotu klūdu novēršana var ievērojami uzlabot datu kvalitāti. Kā daļa no DataOps ir jāievieš automatizētas kontroles un validācijas mehānismi, kas uzrauga datu pareizību, analizē klūdas un identificē modeļus, kā arī fiksē un novērš sistēmas klūdas katrā darba plūsmā. Jo augstāka ir automatizētās validācijas pakāpe, jo augstāka ir kopējā datu kvalitāte un jo mazāka klūdu iespējamība pēdējos posmos.
3. **Dati ir jāpārbauda tāpat kā programmas kods:** lielākā daļa ēku lietojumprogrammu ir balstītas uz datu apstrādi, bet to kontrole bieži vien ir atstāta sekundāru lomu ziņā. Ja mašīnmācīšanās modeļi tiek apmācīti ar neprecīziem datiem, tas noved pie nepareizām prognozēm un finansiāliem zaudējumiem. DataOps ietvaros dati ir jāpākļauj tikpat stingrai kontrolei kā programmatūras kods: logikas pārbaudes, stresa testi, modeļu uzvedības novērtēšana, mainoties ievadvērtībām. Tikai validētus un uzticamus datus var izmantot par pamatu vadības lēmumu pieņemšanai.
4. **Datu novērojamība, neapdraudot veiktpēju:** datu uzraudzība nav tikai metriku apkopojums, bet gan stratēģisks kvalitātes pārvaldības rīks. Lai DataOps darbotos efektīvi, novērojamība ir jāiekļauj visos datu apstrādes posmos, sākot ar projektēšanu un beidzot ar darbību. Tajā pašā laikā ir svarīgi, lai uzraudzība neaizkavētu sistēmas darbību. Būvniecības projektos ir ļoti svarīgi ne tikai vākt datus, bet arī darīt to tā, lai nekādā veidā netiktu traucēts speciālistu

(piemēram, projektētāju), kas veido datus, darbs. Šis līdzsvars ļauj kontrolēt datu kvalitāti, neapdraudot produktivitāti.

DataOps nav papildu slogs datu zinātniekiem, bet gan viņu darba pamatā. Ieviešot DataOps, būvniecības uzņēmumi var pāriet no haotiskas datu pārvaldības uz efektīvu ekosistēmu, kurā dati strādā biznesa labā.

Savukārt VectorOps ir nākamais posms DataOps attīstībā, kas vērts uz daudzdimensiju vektoru datu apstrādi, glabāšanu un analīzi (kas tika aplūkoti iepriekšējā nodalā). Tas ir īpaši svarīgi tādās jomās kā digitālie dvīni, neuronu tīklu modeļi un semantiskā meklēšana, kas sāk ienākt būvniecības nozarē. VectorOps balstās uz vektoru datubāzēm, lai efektīvi uzglabātu, indeksētu un meklētu objektu daudzdimensiju reprezentācijas.

**VectorOps** ir nākamais posms pēc DataOps, kas veltīts vektoru datu apstrādei, analīzei un izmantošanai būvniecībā. Atšķirībā no DataOps, kas koncentrējas uz datu plūsmu, konsekvenci un kvalitāti, VectorOps koncentrējas uz mašīnmācībai nepieciešamo daudzdimensiju objektu reprezentāciju pārvaldību.

Atšķirībā no tradicionālajām pieejām VectorOps ļauj iegūt precīzākus objektu aprakstus, kas ir ļoti svarīgi digitālajiem dvīniem, ģeneratīvās projektēšanas sistēmām un automātiskai kļūdu noteikšanai CAD datos, kas pārveidotī vektoru formātā. Kombinētā DataOps un VectorOps ieviešana veido stabili pamatu mērogojamam, automatizētam darbam ar lieliem informācijas apjomiem - no klasiskām tabulām līdz semantiski bagātīgiem telpiskajiem modeļiem

## Nākamie soļi: no haotiskas glabāšanas uz strukturētu glabāšanu

Tradicionālās pieejas datu uzglabāšanas veidošanai bieži vien rada atšķirīgas "informācijas silosus", kuros svarīgas atziņas nav pieejamas analīzei un lēmumu pieņemšanai. Mūsdienu datu glabāšanas koncepcijas, piemēram, datu krātuve, datu ezers un to hibrīdi, ļauj apvienot atšķirīgu informāciju un padarīt to centralizēti pieejamu datu plūsmas un biznesa izlūkošanas vajadzībām. Ir svarīgi ne tikai izvēlēties pareizo datu glabāšanas arhitektūru, bet arī ieviest datu pārvaldību) un datu minimālismu), lai datu glabātavas nekļūtu par nekontrolējamiem datu purviem).

Apkopojoš šo daļu, ir vērts izcelt galvenos praktiskos soļus, kas palīdzēs jums izmantot aplūkotos jēdzienus ikdienas uzdevumos:

- Izvēlieties efektīvus datu glabāšanas formātus
  - Pāreja no CSV un XLSX uz efektīvākiem formātiem (Apache Parquet, ORC), lai uzglabātu lielu datu apjomu.
  - Datu versiju sistēmas ieviešana, lai izsekotu izmaiņas
  - Metadatu izmantošana, lai aprakstītu informācijas struktūru un izcelsmi.
- Izveidot vienotu uzņēmuma datu arhitektūru
  - Salīdzināt dažādas glabāšanas arhitektūras: RDBMS, DWH un datu ezers. Izvēlieties to,

kas vislabāk atbilst jūsu mērogojamības, avotu integrācijas un analītiskās apstrādes vajadzībām.

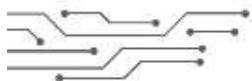
- Izstrādājiet procesa karti datu ieguvei, ielādei un pārveidošanai (ETL) no dažādiem avotiem, lai veiktu savus uzdevumus. Izmantojiet vizualizācijas rīkus, piemēram, Miro, Lucidchart vai Draw.io, lai vizualizētu galvenos soļus un integrācijas punktus.

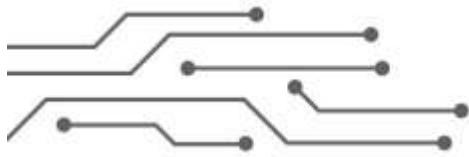
#### ■ Ūstenot datu pārvaldības praksi un datu minimālismu

- Ievērojiet datu minimālisma pieeju - uzglabājiet un apstrādājiet tikai to, kas ir patiesi vērtīgs.
- Ieviest datu pārvaldības principus - noteikt atbildību par datiem, nodrošināt kvalitāti un pārredzamību.
- Uzziniet vairāk par datu pārvaldības politiku un DataOps koncepcijām, VectorOps.
- Datu kvalitātes kritēriju noteikšana un datu validācijas procedūras DataOps ietvaros.

Labi organizēta datu glabāšana rada pamatu uzņēmuma analītisko procesu centralizēšanai. Pāreja no haotiskas failu uzkrāšanas uz strukturētām krātuvēm ļauj pārvērst informāciju par stratēģisku aktīvu, kas palīdz pieņemt pamatotus lēmumus un uzlabot uzņēmējdarbības procesu efektivitāti.

Kad datu vākšanas, pārveidošanas, analīzes un strukturētas uzglabāšanas procesi ir automatizēti un standartizēti, nākamais digitālās transformācijas posms ir pilnvērtīga lielo datu apstrāde.





## IX DAĻA

# LIELIE DATI, MAŠĪNMĀCĪŠANĀS UN PROGNOZES

Devītajā daļā uzmanība pievērsta lielajiem datiem, mašīnmācībai un prognozēšanas analīzei būvniecības nozarē. Tajā aplūkota pāreja no intuitīvas lēmumu pieņemšanas uz objektīvu analīzi, kas balstīta uz vēsturiskiem datiem. Tieki izmantoti praktiski piemēri, lai demonstrētu lielo datu analīzi būvniecībā - no Sanfrancisko būvatlauju datu kopas analīzes līdz CAD apstrādei - projektiem ar miljoniem elementu. Šāda uzmanība tiek pievērsta mašīnmācīšanās metodēm būvniecības projektu izmaksu un termiņu prognozēšanai, detalizēti aplūkojot lineārās regresijas un k-tuvāko kaimiņu algoritmus. Tieki parādīts, kā strukturēti dati kļūst par pamatu prognozēšanas modeļiem, lai novērtētu riskus, optimizētu resursus un uzlabotu projektu vadības efektivitāti. Daļā sniegti arī ieteikumi, kā izvēlēties reprezentatīvus datu paraugus, un paskaidrots, kāpēc efektīvai analīzei ne vienmēr ir nepieciešamas lielas datu kopas.

## NODĀLA 9.1.

### LIELIE DATI UN TO ANALĪZE

#### **Lielie dati būvniecībā: no intuīcijas līdz paredzamībai**

Terminam "lielie dati" nav precīzas definīcijas. Sākotnēji šis jēdziens parādījās, kad informācijas apjoms sāka pārsniegt tradicionālo tās apstrādes metožu iespējas. Mūsdienās datu apjoms un sarežģītība daudzās nozarēs, tostarp būvniecībā, ir palielinājusies tik ļoti, ka tie vairs neiekļaujas datoru lokālajā atmiņā un to apstrādei ir jāizmanto jaunas tehnoloģijas.

Darba ar lielajiem datiem būtība ir ne tikai uzglabāšana un apstrāde, bet arī prognozēšanas iespējas. Būvniecības nozarē lielie dati paver ceļu no intuitīviem lēmumiem, kas balstīti uz tabulu un vizualizāciju subjektīvu interpretāciju (kā minēts iepriekš), uz pamatotām prognozēm, kas balstītas uz reāliem novērojumiem un statistiku.

Pretēji izplatītajam uzskatam, darba ar lielajiem datiem mērķis nav "likt mašīnai domāt kā cilvēkam", bet gan izmantot matemātiskus modeļus un algoritmus, lai analizētu milzīgus datu apjomus, tādējādi identificējot likumsakarības, prognozējot notikumus un optimizējot procesus.

**Lielie dati** nav auksta algoritmu pasaule bez cilvēka ietekmes. Gluži pretēji, lielie dati darbojas kopā ar mūsu instinktiem, kļūdām un radošumu. Tieši cilvēciskās domāšanas nepilnības ļauj mums atrast nestandarta risinājumus un veikt izrāvienu.

Attīstoties digitālajām tehnoloģijām, būvniecības nozare ir sākusi aktīvi izmantot datu apstrādes paņēmienus, kas nāk no IT nozares. Pateicoties tādiem rīkiem kā Pandas un Apache Parquet, var apvienot strukturētus un nestukturētus datus, vienkāršojot piekļuvi informācijai un samazinot analīzes zudumus, savukārt lielas datu kopas no dokumentiem vai CAD projektiem ( 9.2-10. attēls - 9.2-12. attēls ) ļauj apkopot, analizēt un prognozēt datus visos projekta dzīves cikla posmos.

Lielajiem datiem ir pārveidojoša ietekme uz būvniecības nozari, potenciāli ietekmējot to dažādos veidos. Lielo datu tehnoloģiju izmantošana dod rezultātus vairākās svarīgās jomās, tostarp, piemēram, šādās:

- **Investīciju potenciāla analīze** - projektu rentabilitātes un atmaksāšanās perioda prognozēšana, pamatojoties uz iepriekšējo objektu datiem.
- **Paredzamā apkope** - iespējamu aprīkojuma kļūmju identificēšana, pirms tās faktiski rodas, tādējādi samazinot dīkstāves laiku.
- **Piegādes ķēdes optimizācija** - traucējumu prognozēšana un loģistikas efektivitātes uzlabošana.
- **Energoefektivitātes analīze** - palīdzība zema enerģijas patēriņa ēku projektēšanā.
- **Drošības uzraudzība** - sensoru un valkājamu ierīču izmantošana, lai uzraudzītu darba vietas apstākļus.

- **Kvalitātes kontrole** - reāllaika uzraudzība attiecībā uz atbilstību procesa standartiem.
- **Darbaspēka pārvaldība** - darba rezultātu analīze un personāla vajadzību prognozēšana.

Būvniecībā ir grūti atrast jomu, kurā datu analītika un prognozes nebūtu pieprasītas. Galvenā prognozēšanas algoritmu priekšrocība ir to spēja pašmācīties un pastāvīgi pilnveidoties, uzkrājot datus.

Tuvākajā nākotnē mākslīgais intelekts ne tikai palīdzēs celtniekim, bet arī pieņems svarīgākos lēmumus - sākot ar projektēšanas procesiem un beidzot ar ēku ekspluatācijas jautājumiem.

Plašāk par to, kā tiek ģenerēti prognozes un izmantoti mācīšanās modeļi, tiks aplūkots nākamajā grāmatas daļā "Mašīnmācīšanās un prognozes".

Pārejai uz pilnvērtīgu darbu ar lielajiem datiem ir jāmaina pati pieeja analītikai. Ja līdz šim aplūkotās klasiskās sistēmas koncentrējās uz cēloņu un seku sakarībām, tad lielo datu analītika pārorientējas uz statistisko modeļu un korelāciju meklēšanu, kas ļauj noteikt slēptās sakarības un prognozēt objektu uzvedību pat bez pilnīgas izpratnes par visiem faktoriem.

### **Jautājums par lielo datu izmantošanas iespējamību: korelācija, statistika un datu izlase**

Tradicionāli konstruēšana balstījās uz subjektīvām hipotēzēm un personīgo pieredzi. Inženieri ar zināmu varbūtības pakāpi pieņēma, kā uzvedīsies materiāls, kādas slodzes izturēs konstrukcija un cik ilgi projekts kalpos. Šie pieņēmumi tika pārbaudīti praksē, bieži vien uz laika, resursu un nākotnes risku rēķina.

Līdz ar lielo datu apjomu parādīšanos pieeja būtiski mainās: lēmumi vairs netiek pieņemti, pamatojoties uz intuitīvām nojausmām, bet gan analizējot liela apjoma datu kopas.

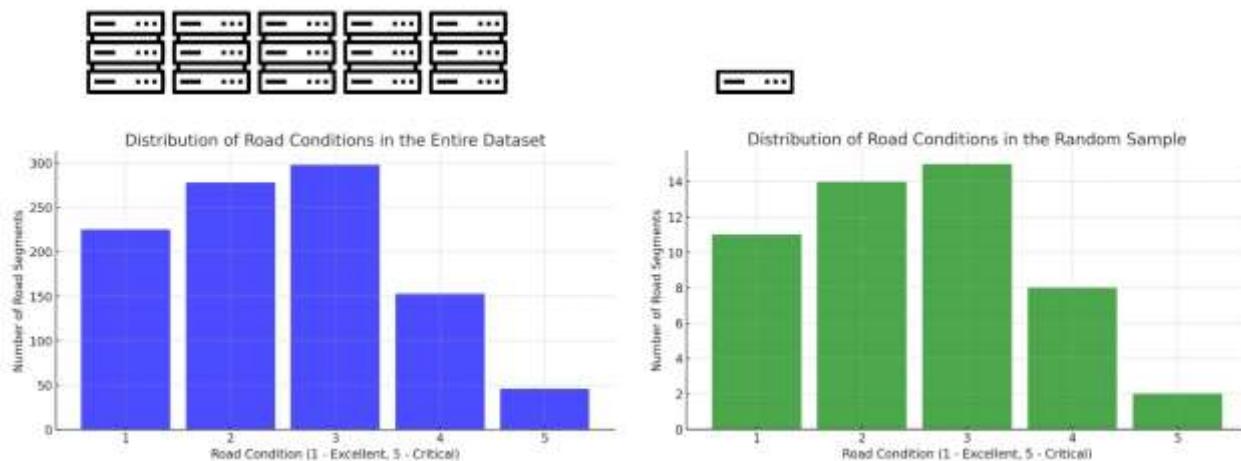
Būvniecība pamazām pārstāj būt intuīcijas māksla un kļūst par precīzu prognozēšanas zinātni.

Pāreja uz ideju par lielo datu izmantošanu neizbēgami rada svarīgu jautājumu: cik būtisks ir datu apjoms un cik daudz informācijas patiešām ir nepieciešams, lai veiktu uzticamu prognozēšanas analīzi? Plaši izplatītais uzskats, ka "jo vairāk datu, jo lielāka precīzitāte", praksē ne vienmēr izrādās statistiski pamatots.

Jau 1934. gadā statistikis Jeržijs Neimans pierādīja [145], ka statistisko secinājumu precīzitātes atslēga ir ne tik daudz datu apjomā, cik to reprezentativitātē un izlases veida nejaušībā.

Tas jo īpaši attiecas uz būvniecības nozari, kur tiek vākti lieli datu masīvi, izmantojot IoT - sensorus, skenerus, novērošanas kameras, bezpilota lidaparātus un pat daudzformāta CAD - modeļus, kas palielina "aklo punktu", noviržu un datu izkroplojumu risku.

Aplūkosim piemēru par ceļa seguma stāvokļa uzraudzību. Pilns visu ceļa posmu datu kopums var aizņemt X GB, un tā apstrādei var būt nepieciešama aptuveni diena. Tajā pašā laikā izlases paraugam, kurā iekļauts tikai katrs 50. ceļa posms, būtu nepieciešams tikai  $X/50$  GB, un to apstrādātu pusstundas laikā, vienlaikus nodrošinot līdzīgu aplēšu precizitāti noteiktiem aprēķiniem (9.1-1. attēls).



9.1-1. attēls Bruža seguma stāvokļa histogrammas: pilns datu kopums un izlases veida izlase uzrāda identiskus rezultātus.

Tādējādi veiksmīgas datu analīzes atslēga bieži vien var būt nevis datu apjoms, bet gan parauga reprezentativitāte un izmantoto apstrādes metožu kvalitāte. Pārejai uz izlases veida izlasi un selektīvāku pieeju ir nepieciešama domāšanas maiņa būvniecības nozarē. Vēsturiski uzņēmumi ir vadījušies pēc loģikas "jo vairāk datu, jo labāk", uzskatot, ka, aptverot visus iespējamos rādītājus, tiks panākta maksimāla precizitāte.

Šāda pieeja atgādina populāru maldīgu priekšstatu par projektu vadību: "Jo vairāk speciālistu piesaistīšu, jo efektīvāks būs darbs". Tomēr, tāpat kā cilvēkresursu gadījumā, svarīgāka ir kvalitāte un instrumenti, nevis kvantitāte. Nejemanot vērā datu vai projekta dalībnieku savstarpējās sakarības (korelācijas), apjoma palielināšana var radīt tikai troksni, izkroplojumus, dublēšanos un nevajadzīgu izšķērdēšanu.

Galu galā bieži vien izrādās, ka daudz produktīvāk ir izmantot mazāku, bet kvalitatīvi sagatavotu datu kopu, kas spēj sniegt stabilas un pamatotas prognozes, nekā paļauties uz milzīgu, bet haotisku informāciju, kurā ir daudz pretrunīgu signālu.

Pārmērīgs datu apjoms ne tikai negarantē lielāku precizitāti, bet arī var novest pie izkroplojotiem secinājumiem trokšņu, lieku pazīmju, slēptu korelāciju un nebūtiskas informācijas klātbūtnes dēļ. Šādos apstākļos palielinās pārmērīgas modeļu pielāgošanas risks un samazinās analītisko rezultātu ticamība.

Būvniecībā galvenais izaicinājums, strādājot ar lielajiem datiem, ir noteikt optimālo datu daudzumu un kvalitāti. Piemēram, uzraugot betona konstrukciju stāvokli, izmantojot tūkstošiem sensoru un vācot informāciju katru minūti, datu uzglabāšanas un analīzes sistēma var būt pārslogota. Tomēr, veicot

korelācijas analīzi un atlasot 10% informatīvāko sensoru, var iegūt gandrīz identisku prognožu precizitāti, iztērējot daudzkārt, dažkārt desmitiem un simtiem reižu mazāk resursu.

Izmantojot mazāku datu apakškopu, samazinās gan nepieciešamās datu glabāšanas apjoms, gan apstrādes laiks, kas ievērojami samazina datu glabāšanas un analīzes izmaksas un bieži vien padara izlases veida izlasi par ideālu risinājumu prognozēšanas analītikai, jo īpaši lielos infrastruktūras projektos vai strādājot reālajā laikā. Galu galā būvniecības procesu efektivitāti nosaka nevis savākto datu apjoms, bet gan to analīzes kvalitāte. Bez kritiskas pieejas un rūpīgas analīzes dati var novest pie nepareiziem secinājumiem.

Pēc noteikta datu apjoma katra jauna informācijas vienība dod arvien mazāk un mazāk noderīgu rezultātu. Tā vietā, lai bezgalīgi vāktu informāciju, ir svarīgi pievērst uzmanību tās reprezentativitātei un analīzes metodēm ( 9.2-2. attēls).

Šo parādību labi aprakstījis Allen Wallis [146], kurš ilustrē statistikas metožu izmantošanu, izmantojot piemēru ar divu alternatīvu ASV Jūras kara flotes šāviņu konstrukciju testēšanu.

*Jūras kara flote pārbaudīja divas alternatīvas šāviņu konstrukcijas (A un B), veicot virknī apšaudījumu pa pāriem. Katrā raundā A saņem 1 vai 0 atkarībā no tā, vai tā veikspēja ir labāka vai sliktāka nekā B, un otrādi. Standarta statistiskā pieeja ietver fiksēta izmēģinājumu skaita (piemēram, 1000) veikšanu un uzvarētāja noteikšanu, pamatojoties uz procentuālo sadalījumu (piemēram, ja A saņem 1 vairāk nekā 53% gadījumu, tas tiek uzskatīts par labāko). Kad Allens Voliss apsprieda šādu problēmu ar (Jūras spēku) kapteinī Garetu L. Šuileru (Garrett L. Schuyler), kapteinis iebilda, ka šāds tests, citējot Allena teikto, varētu būt bezjēdzīgs. Ja uz vietas būtu bijis tāds gudrs un pieredzējis munīcijas virsnieks kā Šuilers, viņš pēc pirmajiem dažiem simtiem [šāvienu] būtu redzējis, ka eksperiments nav jāpārtrauc vai nu tāpēc, ka jaunā metode ir acīmredzami sliktāka, vai arī tāpēc, ka tā ir acīmredzami labāka par to, uz ko tika cerēts [146].*

- ASV valdības Statistikas pētījumu grupa Kolumbijas Universitātē, Otrā pasaules kara periods.

Šo principu plaši izmanto dažādās nozarēs. Piemēram, medicīnā jaunu zāļu klīniskie pētījumi tiek veikti ar nejauši atlaistiem pacientu paraugiem, kas ļauj iegūt statistiski nozīmīgus rezultātus, netestējot zāles ar visiem planētas iedzīvotājiem. Ekonomikā un socioloģijā tiek veiktas reprezentatīvas aptaujas, lai atspoguļotu sabiedrības viedokli, neaptaujājot visus valsts iedzīvotājus.

Līdzīgi kā valdības un pētniecības organizācijas veic nelielu iedzīvotāju grupu aptaujas, lai izprastu vispārējās sociālās tendences, arī būvniecības nozares uzņēmumi var izmantot izlases datu paraugus, lai efektīvi uzraudzītu un veidotu prognozes projektu vadībai ( 9.1-1. attēls).

*Lielie dati var mainīt pieeju sociālajām zinātnēm, taču tie neaizstās statistisko veselo saprātu [147].*

- Thomas Landsall-Welfair, "Forecasting the nation's current mood", Significance v. 9(4), 2012 .

No resursu taupīšanas viedokļa, vācot datus nākotnes prognozēm un lēmumu pieņemšanai, ir svarīgi atbildēt uz jautājumu: vai ir jēga tērēt ievērojamus resursus milzīgu datu kopu vākšanai un apstrādei, ja var izmantot daudz mazāku un lētāku testa datu kopu, ko var pakāpeniski palielināt? Nejaušās izlases efektivitāte liecina, ka uzņēmumi var desmitiem vai pat tūkstošiem reižu samazināt izmaksas par modeļu vākšanu un apmācību, izvēloties datu vākšanas metodes, kas neprasā visaptverošu pārkājumu, bet joprojām nodrošina pietiekamu precizitāti un reprezentativitāti. Šāda pieeja ļauj pat maziem uzņēmumiem, izmantojot ievērojami mazāk resursu un datu apjomu, sasniegt rezultātus, kas ir līdzvērtīgi lielajām korporācijām, un tas ir svarīgi uzņēmumiem, kuri vēlas optimizēt izmaksas un paātrināt pamatotu lēmumu pieņemšanu, izmantojot nelielus resursus. Turpmākajās nodaļās izpētiet analītikas un prognozēšanas analītikas piemērus, kas balstīti uz publiskām datu kopām, izmantojot lielo datu rīkus.

## **Lielie dati: datu analīze no Sanfrancisko miljoniem būvatļauju datu kopas**

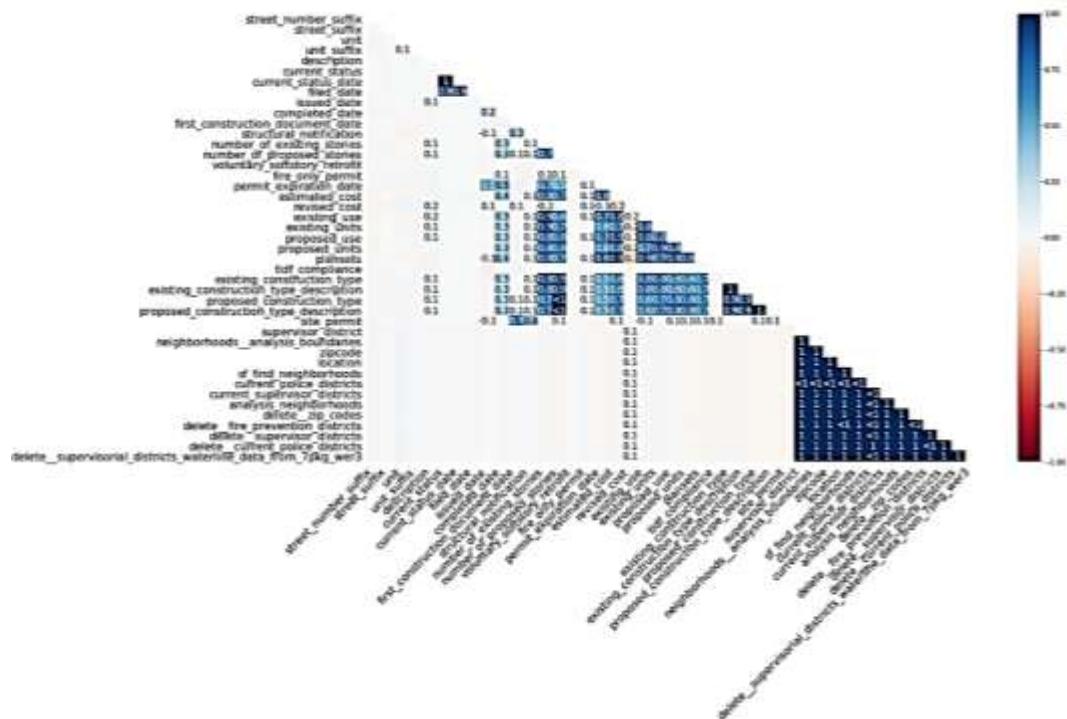
Darbs ar atvērtām datu kopām sniedz unikālu iespēju praktiski pielietot iepriekšējās nodaļās aplūkotos principus: pārdomātu pazīmju atlasi, reprezentatīvu izlasi, vizualizāciju un kritisku analīzi. Šajā nodaļā mēs pētīsim, kā, izmantojot atvērtos datus, var izpētīt tādas sarežģītas parādības kā būvniecības aktivitātes lielā pilsētā - jo īpaši vairāk nekā miljons ierakstu par būvniecības atļaujām Sanfrancisko.

Publiski pieejami dati par vairāk nekā vienu miljonu būvatļauju (attēls 9.1-2) (ieraksti divās datu kopās CSV formātā) no "Sanfrancisko Būvniecības departamenta". [148] ļauj mums izmantot neapstrādāto CSV -tabulu, lai analizētu ne tikai būvniecības aktivitāti pilsētā,, bet arī kritiski analizētu nesenā tendenci un Sanfrancisko būvniecības nozares vēsturi pēdējo 40 gadu laikā, no 1980. līdz 2019. gadam.

Datu kopas vizualizāciju izveidei izmantotie koda piemēri (attēli 9.1-3- attēli 9.1-8), kā arī vizuālie grafiki ar kodu, paskaidrojumiem un komentāriem ir atrodami Kaggle platformā, meklējot "San Francisco. Būvniecības nozare 1980-2019." [149].

count 1.137695e+06		Building Permits on or after January 1, 2013			Building Permits before January 1, 2013		
permit_creation_date	description	current_status	current_status_date	filed_date	issued_date	completed_date	
07/01/1998	repair stucco	complete	07/07/1998	07/01/1998	07/01/1998	07/07/1998	
12/13/2004	reroofing	expired	01/24/2006	12/13/2004	12/13/2004	NaN	
02/18/1992	install auto fire spks.	complete	06/29/1992	02/18/1992	03/18/1992	06/29/1992	
permit_number	permit_expiration_date	estimated_cost	revised_cost	existing_use	Zipcode	Location	
362780	9812394	11/01/1998	780.0	NaN	1 family dwelling	94123.0 (37.796246876498, -122.4322541443574)	
570817	200412131233	06/13/2005	9000.0	9000.0	apartments	94127.0 (37.729258516008388, -122.4644245957462)	
198411	9202396	09/18/1992	9000.0	NaN	apartments	94111.0 (37.795086002552974, -122.39593224461805)	

Attēlā 9.1-2 Datu kopas satur informāciju par izsniegtajām būvatļaujām ar dažādiem objektu atribūtiem.

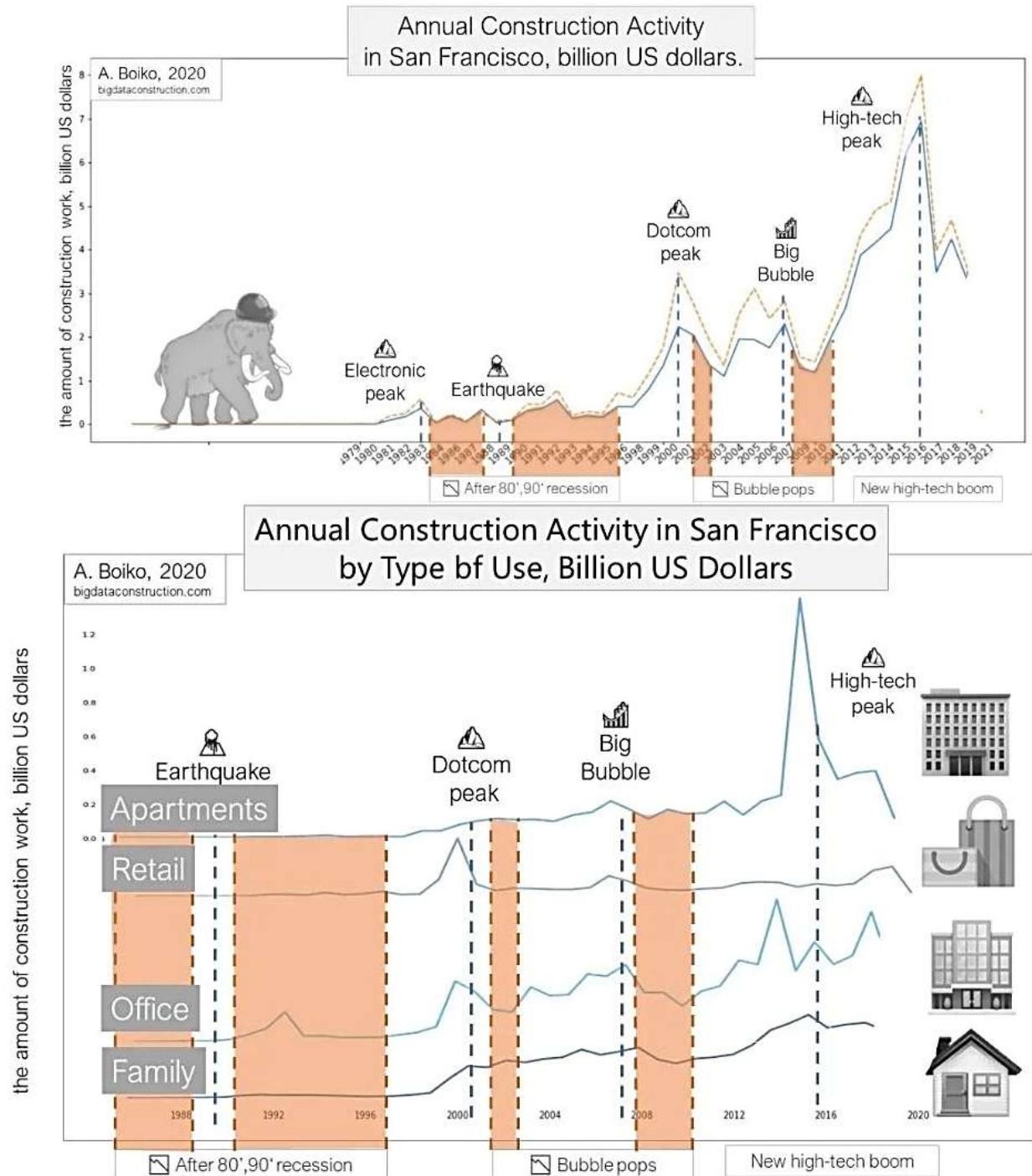


Attēls. 9.1-3 Siltuma karte (Pandas un Seaborn), kas vizualizē visus datu kopas atribūtus un palīdz noteikt attiecības starp atribūtu pāriem.

No Sanfrancisko Būvniecības departamenta iesniegtās tabulas ( 9.1-2. attēls) nav redzamas nekādas tendences vai secinājumi. Sausie skaitļi tabulas formā nav pamats lēmumu pieņemšanai. Lai datus padarītu vizuāli saprotamus, kā detalizēti aprakstīts nodaļās par datu vizualizāciju, tie jāvizualizē, izmantojot dažādas bibliotēkas, kas aplūkotas grāmatas septītajā daļā par tēmu "ETL un rezultātu vizualizēšana grafiku veidā".

Analizējot datus, izmantojot Pandas DataFrame un Python vizualizācijas bibliotēkas<sup>1</sup> 137 695 atļauju vērtību [148], var secināt, ka būvniecības aktivitāte Sanfrancisko ir cieši saistīta ar ekonomiskajiem cikliem, jo īpaši ar strauji augošo Silīcija ielejas tehnoloģiju nozari ( 9.1-4. attēls).

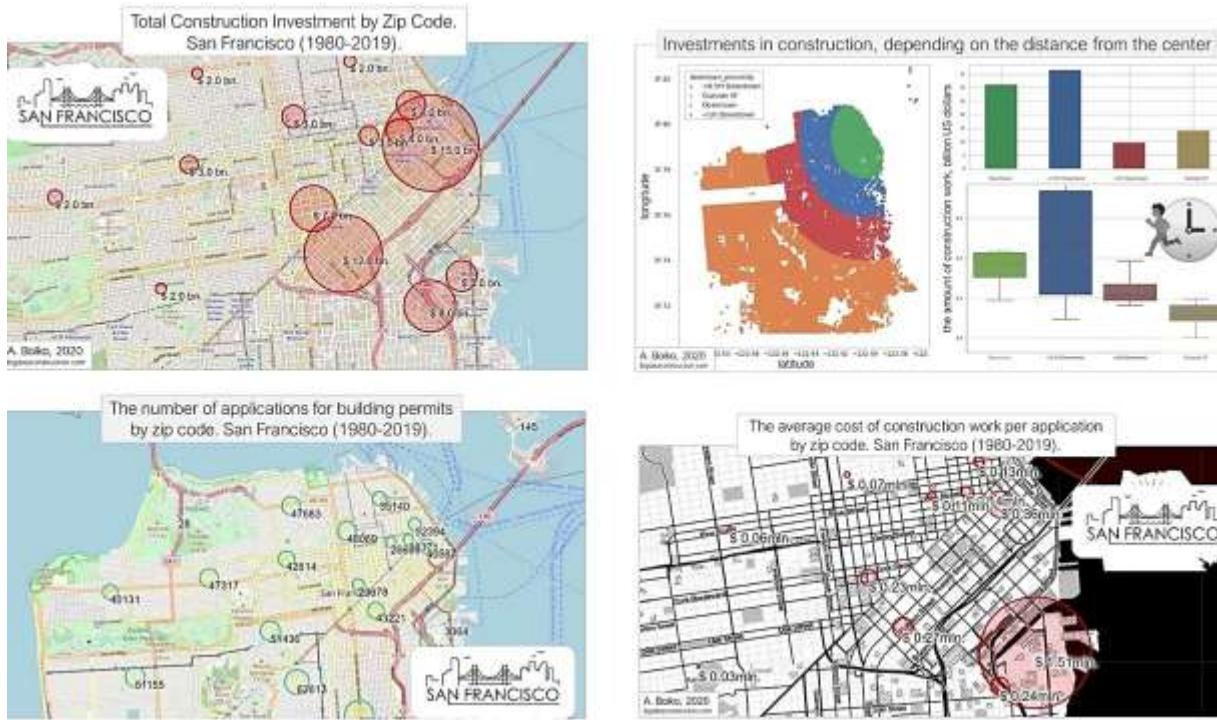
Ekonomikas uzplaukums un lejupslīde būtiski ietekmē būvniecības projektu skaitu un vērtību. Piemēram, pirmais būvniecības aktivitātes maksimums sakrita ar elektronikas bumu 80. gadu vidū (izmantoja Pandas un Matplotlib), un nākamie maksimumi un kritumi bija saistīti ar dot-com burbuli un pēdējo gadu tehnoloģiju bumu.



Attēls 9.1-4 Sanfrancisko nekustamā īpašuma sektorā investīcijas korelē ar Silīcija ielejas tehnoloģisko attīstību.

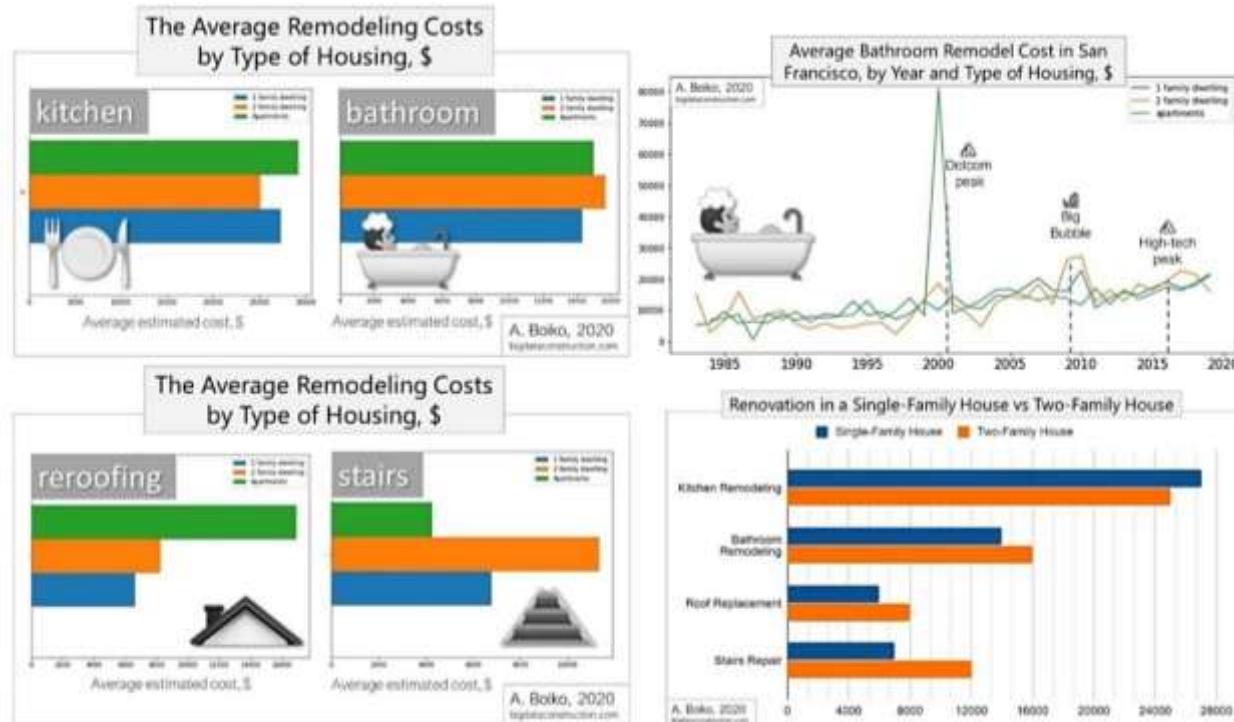
Datu analīze liecina, ka Sanfrancisko lielākā daļa no 91,5 miljardiem ASV dolāru, kas pēdējo desmit gadu laikā ieguldīti būvniecībā un pārbūvē, - gandrīz 75% - ir koncentrēti pilsetas centrālajā daļā (attēls 9.1-5 - izmantota Pandas un Folium vizualizācijas bibliotēka) un 2 km rādiusā no centra, kas atspoguļo lielāku ieguldījumu blīvumu šajās centrālajās zonās.

Vidējās būvatļauju izmaksas ievērojami atšķiras atkarībā no apkaimes, un pilsētas centrā iesniegtie pieteikumi izmaksā trīs reizes dārgāk nekā ārpus pilsētas centra, jo ir lielākas zemes, darbaspēka un materiālu izmaksas, kā arī stingrāki būvniecības noteikumi, kas prasa dārgākus materiālus, lai uzlabotu energoefektivitāti.



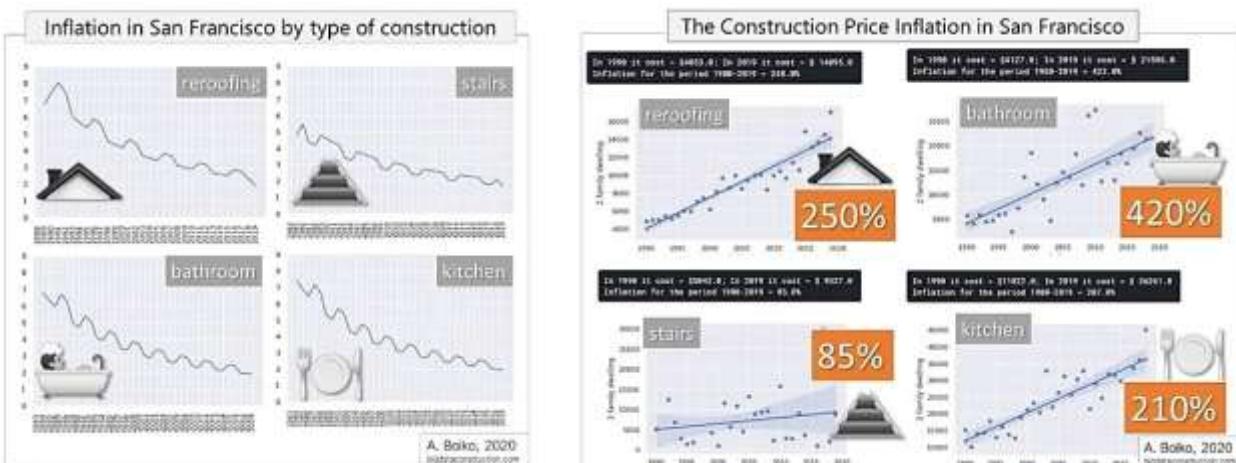
Attēls 9.1-5 Sanfrancisko 75 procenti būvniecības investīciju (91,5 miljardi ASV dolāru) ir koncentrēti pilsētas centrā.

Datu kopa ļauj aprēķināt vidējās remonta cenas ne tikai pēc māju tipa, bet arī pēc pilsētas rajoniem un atsevišķām adresēm (pasta indeksiem). Sanfrancisko mājokļu remonta izmaksu dinamika parāda atšķirīgas tendences dažādiem remonta veidiem un mājokļiem (attēls 9.1-6 izmantoti Pandas un Matplotlib). Virtuves renovācija ir ievērojami dārgāka nekā vannas istabas renovācija: vidēji virtuves renovācija vienģimenes mājā izmaksā aptuveni 28 000 ASV dolāru, bet divģimenes mājā - 25 000 ASV dolāru.



Attēls 9.1-6 SF virtuves remonts izmaksā gandrīz divreiz dārgāk nekā vannas istabas remonts, un mājokļu īpašniekiem 15 gadu laikā katrai mēnesi jāatvēl \$ 350, lai segtu mājokļa kapitālo remontu izmaksas.

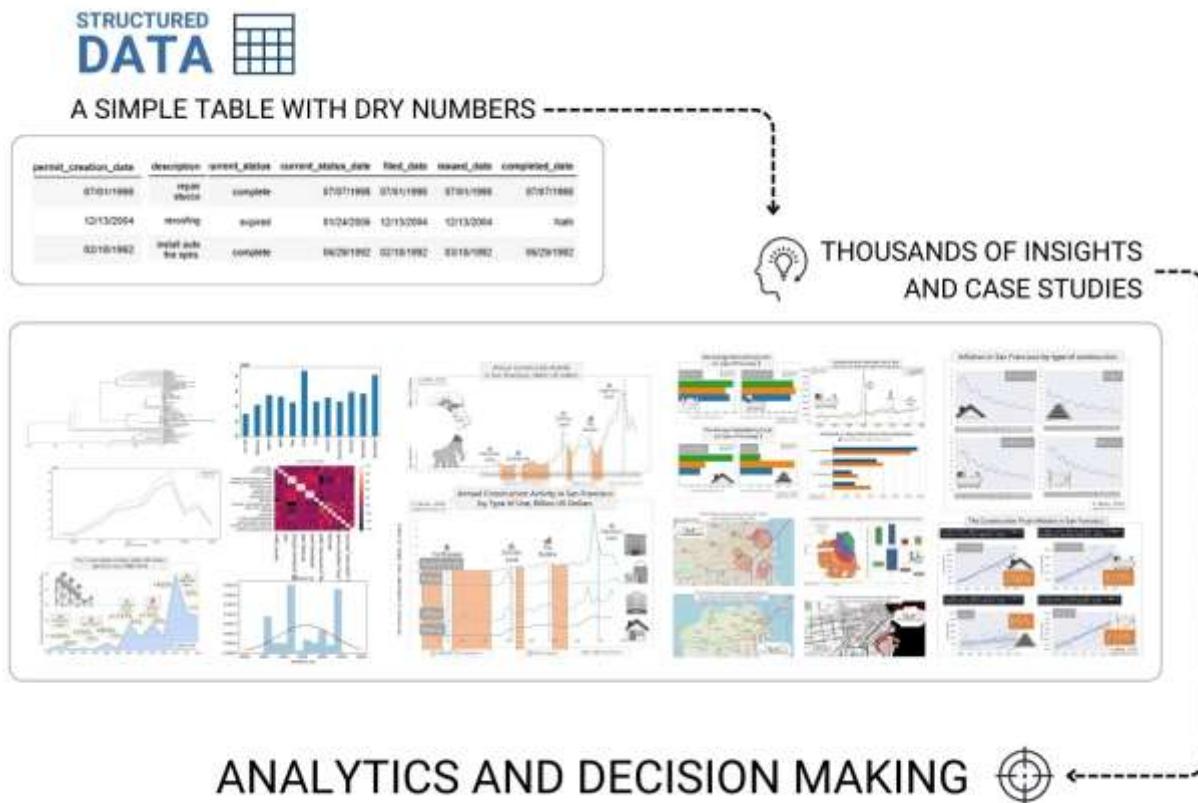
Būvniecības izmaksu inflāciju Sanfrancisko gadu gaitā var izsekot, analizējot datus, kas sagrupēti pa mājokļu tipiem un gadiem (attēls 9.1-7 - izmantots Pandas un Seaborn), kurā redzams stabils vidējo remonta izmaksu pieaugums kopš 1990. gada un atklāti īstermiņa trīs gadu cikli daudzdzīvokļu ēku remonta izmaksās.



Attēls 9.1-7 No 1980. līdz 2019. gadam vannas istabas renovācijas izmaksas SF ir pieckāršojušās, savukārt jumta un virtuves renovācijas izmaksas ir trīskāršojušās, bet kāpņu telpu renovācijas izmaksas ir pieaugašas tikai par 85%.

Sanfrancisko Būvniecības departamenta publisko datu izpēte (9.1-3. attēls) atklāj, ka būvniecības izmaksas pilsētā ir ļoti mainīgas un bieži neparedzamas, un tās ietekmē dažādi faktori. Šie faktori ietver ekonomisko izaugsmi, tehnoloģiskās inovācijas un dažādu mājokļu tipu unikālās prasības.

Agrāk šādām analīzēm bija nepieciešamas padzīlinātas zināšanas programmēšanā un analīzē. Tomēr līdz ar LLM- rīku parādīšanos šis process ir kļuvis pieejams un saprotams plašam būvniecības nozares profesionāļu lokam, sākot no inženieriem projektēšanas nodaļās un beidzot ar augstāko vadību.



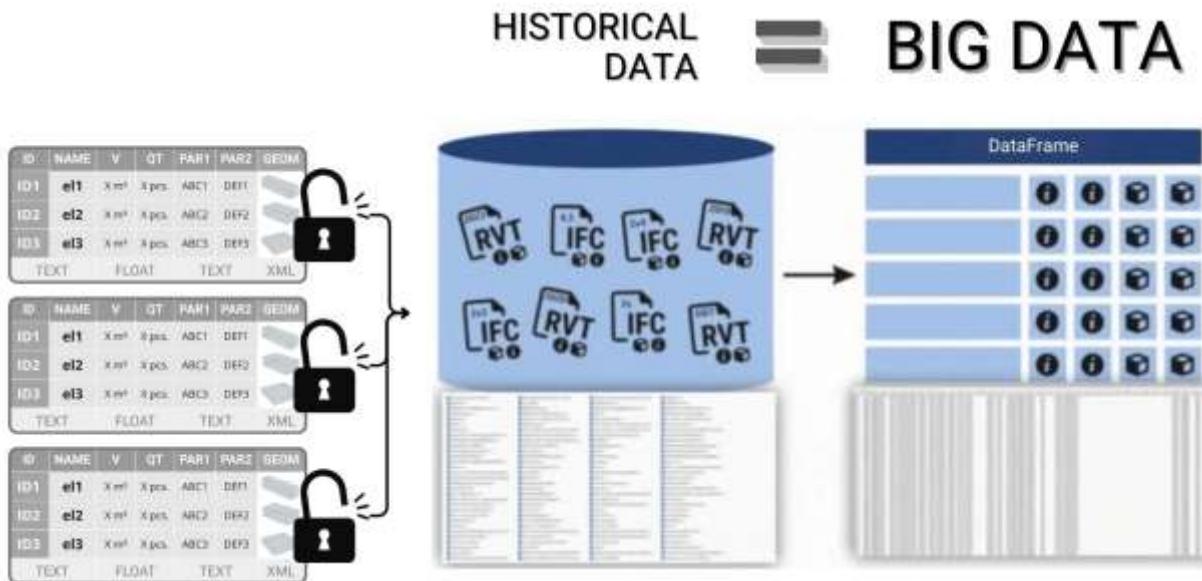
Attēls 9.1-8 Pāreja uz vizuāli saprotamiem datiem īauj automatizēti pieņemt lēmumus, atpazīstot slēptos modeļus.

Tāpat kā mēs analizējām datus no "Sanfrancisko Būvniecības pārvaldes" tabulas datu kopas, mēs varam vizualizēt un analizēt jebkuru datu kopu - no attēliem un dokumentiem līdz IoT datiem vai datiem no atvasinātām CAD datubāzēm.

## Lielo datu piemērs, pamatojoties uz CAD datiem (BIM)

Šajā piemērā mēs analizēsim lielu datu kopu, izmantojot datus no dažādiem CAD rīkiem (BIM). Lai savāktu un izveidotu lielo datu kopu, tika izmantots specializēts automatizēts tīmekļa pārlūkošanas rīks (skripts), kas konfigurēts automātiskai meklēšanai un projektēšanas failu vākšanai no tīmekļa vietnēm, kurās tiek piedāvāti bezmaksas arhitektūras modeļi RVT un IFC formātos. Dažu dienu laikā pārlūks veiksmīgi atrada un lejupielādēja 4596 IFC failus, 6471 RVT failu un 156 024 DWG failus[149].

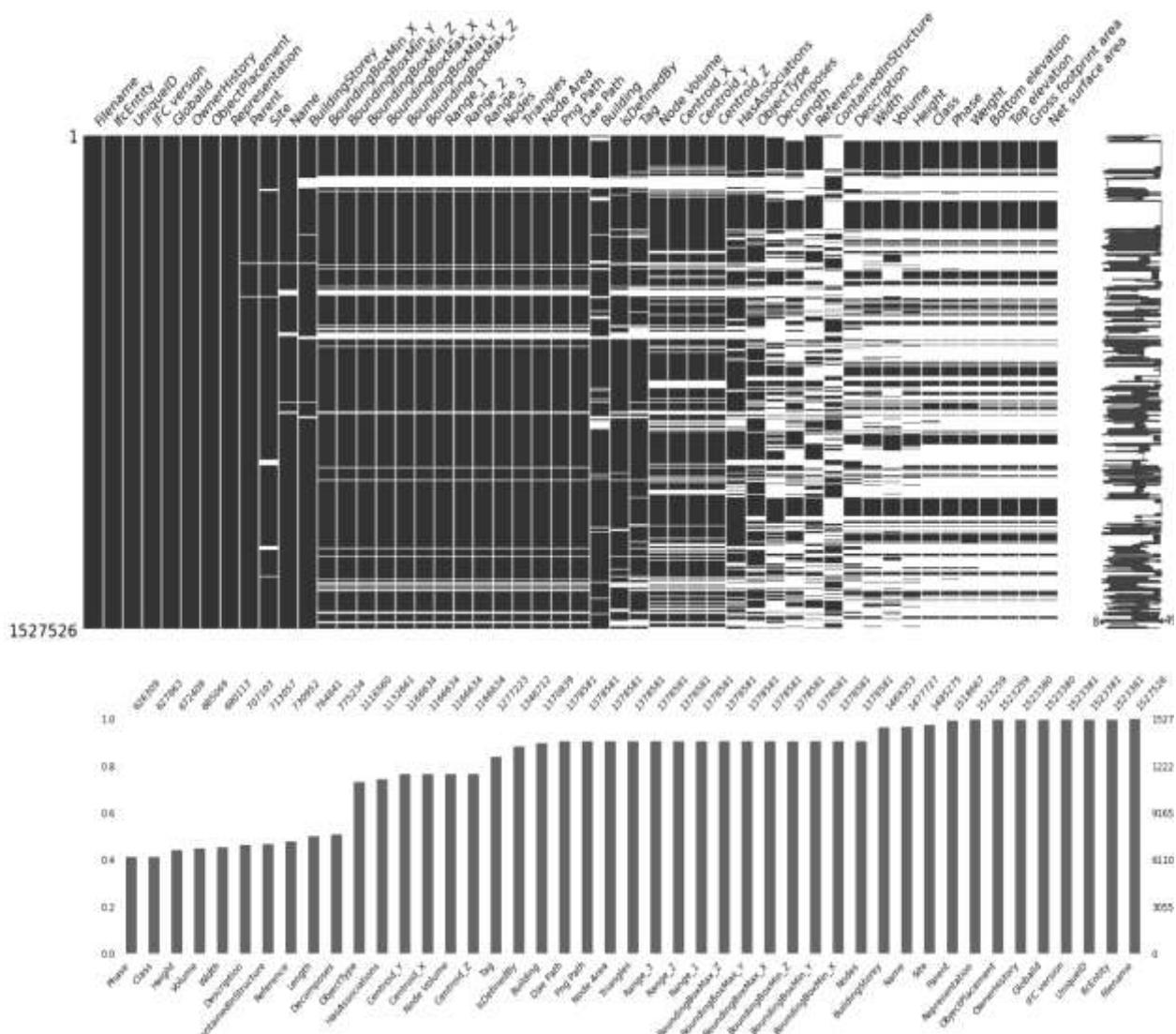
Pēc projektu apkopošanas dažādu versiju RVT un IFC formātos un to konvertēšanas strukturētā CSV formātā, izmantojot bezmaksas reversās inženierijas SDK, gandrīz 10 tūkstoši RVT un IFC projektu tika apkopoti vienā lielā Apache Parquet tabulas failā un augšupielādēti Pandas DataFrame analīzei (9.1-9. attēls).



**Attēls 9.1-9 Strukturētie dati projektu dati ļauj apvienot jebkuru projektu skaitu vienā divdimensiju tabulā.**

Šīs plašās kolekcijas dati satur šādu informāciju: IFC datņu kopa satur aptuveni 4 miljonus vienību (rindu) un 24 962 atribūtus (kolonnas), un RVT datņu kopa, kas sastāv no aptuveni 6 miljoniem vienību (rindu), satur 27 025 dažādus atribūtus (kolonnas).

Šīs informācijas kopas (9.1-10. attēls) aptver miljoniem elementu, no kuriem katram tika papildus iegūtas un kopējā tabulā pievienotas Bounding Box ģeometrijas koordinātas (taisnstūris, kas nosaka objekta robežas projektā) - tika izveidotas Bounding Box ģeometrijas koordinātas (taisnstūris, kas nosaka objekta robežas projektā) un katra elementa attēli PNG formātā un ģeometrija atvērtā XML formātā - DAE (Collada).



**Attēls 9.1-10 1,5 miljonu elementu Sabsēt un pirmo 100 atrībūtu aizņemtības vizualizācija (trūkstošāno bibliotēka) histogrammas veidā.**

Tādējādi mēs ieguvām visu informāciju par desmitiem miljonu elementu no 4596 IFC projektiem un 6471 RVT projekta, kur visi visu esību elementu atrībūti-īpašības un to ģeometrija (Bounding Box) tika pārveidoti strukturētā vienas tabulas formā (DataFrame) (9.1-10. attēls - dati par datu rāmju populācijām parādās histogrammu veidā).

Analīzes procesa laikā uzzīmētās histogrammas ( 9.1-10. attēls, 9.2-6. attēls, 9.2-7. attēls) ļauj ātri novērtēt datu blīvumu un vērtību biežumu kolonnās. Tas sniedz pirmo ieskatu par pazīmju sadalījumu, noviržu klātbūtni un atsevišķu pazīmju potenciālo lietderību, analizējot un veidojot mašīnmācīšanās modeļus.

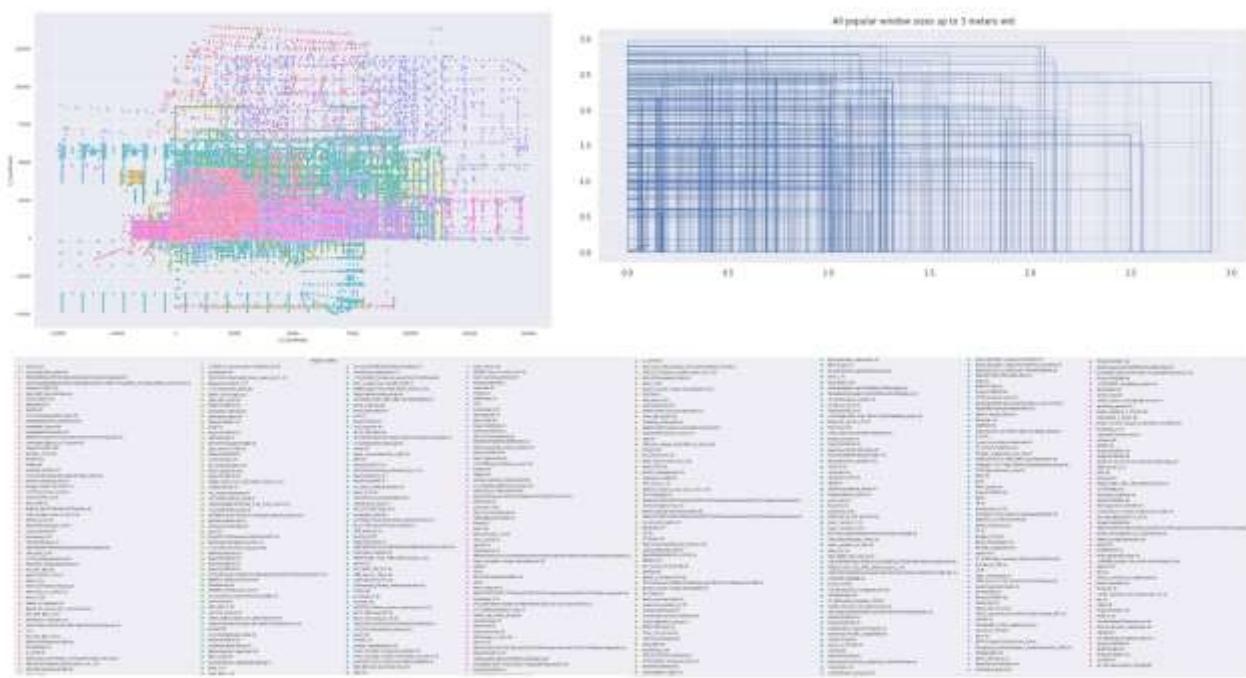
Viens no piemēriem, kā šo datu kopu var izmantot praksē ( 9.1-10. attēls), ir projekts "5000 IFC un RVT ". [149], kas pieejams Kaggle platformā. Tajā ir prezentēta Jupyter piezīmju grāmata ar pilnīgu Pipeline

risinājumu: no datu pirmapstrādes un analīzes līdz rezultātu vizualizācijai, izmantojot Python bibliotēkas - pandas, matplotlib, seaborn, folium un citas (9.1-11. att.).



**Attēls 9.1-11** Datu analīzes piemēri no CAD formātiem (BIM), izmantojot Python vizualizācijas bibliotēkas un pandas bibliotēku.

Pamatojoties uz metainformāciju, ir iespējams noteikt, kurās pilsētās ir izstrādāti konkrēti projekti, un parādīt to kartē (piemēram, izmantojot folium bibliotēku). Turklat laika zīmogi datos ļauj izpētīt datņu saglabāšanas vai rediģēšanas laika modeļus: pēc nedēļas dienas, diennakts laika un mēneša.



**Attēls 9.1-12** Vizualizācija visu kolonnu ģeometriskā novietojuma un visu logu izmēru līdz 3 metriem projektos no saraksta diagrammas apakšā.

No modeļiem iegūtie ģeometriskie parametri Bounding Box formā ir piemēroti arī apkopotai analīzei.

Piemēram, vietnē 9.1-12. attēlā parādīti divi grafiki: kreisajā attēlots attālumu starp kolonnām sadalījums visiem projektiem attiecībā pret nulles punktu, bet labajā attēlots visu logu līdz 3 metru augstumam izmēri desmitiem tūkstošu logu elementu paraugā (pēc visas datu kopas grupēšanas pēc parametra "Category" (kategorija) ar vērtību "OST\_Windows", "IfcWindows").

Šī piemēra Pipeline analīzes kods un pati datu kopa ir pieejama Kaggle tīmekļa vietnē ar nosaukumu "5000 IFC un RVT | DataDrivenConstruction.io projects" [149]. Šo pabeigto cauruļvadu kopā ar datu kopu var kopēt un bez maksas palaist tiešsaistē Kaggle vai bezsaistē kādā no populārākajām IDE: PyCharm, Visual Studio Code (VS Code), Jupyter Notebook, Spyder, Atom, Sublime Text, Eclipse ar PyDev spraudni, Thonny, Wing IDE, IntelliJ IDEA ar Python spraudni, JupyterLab vai populārajiem tiešsaistes rīkiem Kaggle.com, Google Collab, Microsoft Azure Notebooks, Amazon SageMaker.

Analītiskajām atziņām, kas iegūtas, apstrādājot un pētot milzīgus strukturētu datu apjomus, būs izšķiroša nozīme būvniecības nozares lēmumu pieņemšanas procesos.

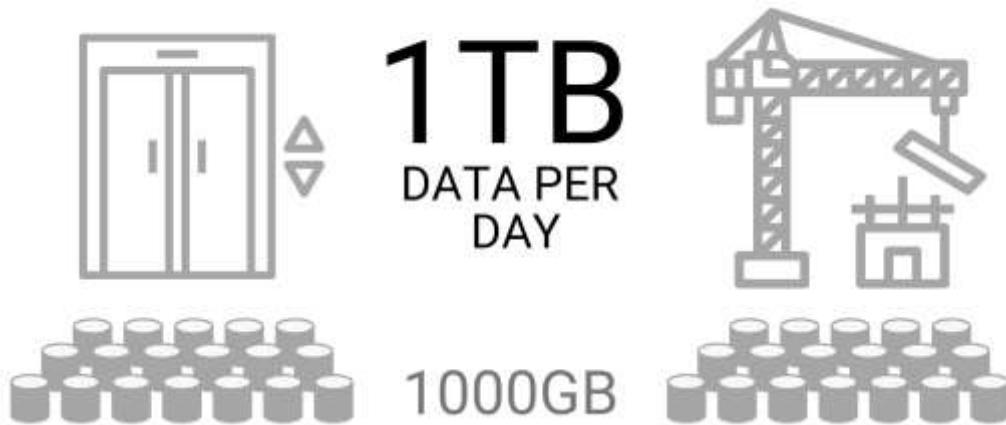
Izmantojot šāda veida informācijas analīzi, kas balstīta uz iepriekšējiem projektiem, speciālisti var efektīvi prognozēt, piemēram, materiālu un darbaspēka vajadzības un optimizēt projektēšanas risinājumus pirms būvniecības uzsākšanas.

Tomēr, ja projektēšanas dati vai būvatļaujas ir relatīvi statiska informācija, kas mainās salīdzinoši lēni, tad pats būvniecības process strauji piesātinās ar dažādiem sensoriem un IoT -ierīcēm: kamerām, automatizētām monitoringa sistēmām, kas pārraida datus reāllaikā - tas viss pārvērš būvlaukumu par dinamisku digitālu vidi, kurā dati jāanalizē reāllaikā.

## **IoT Lietu internets un viedie līgumi**

Lietu internets Lietu internets ir jauns digitālās transformācijas vilnis, kurā katras ierīce saņem savu IP adresi un kļūst par globāla tīkla daļu. IoT ir koncepcija, kas ietver fizisku objektu pieslēgšanu internetam, lai vāktu, apstrādātu un pārsūtītu datus. Būvniecībā tas nozīmē iespēju kontrolēt būvniecības procesus reāllaikā, samazināt materiālu zudumus, prognozēt iekārtu nolietojumu un automatizēt lēmumu pieņemšanu.

Saskaņā ar CFMA rakstu "Gatavojoties nākotnei ar savienoto būvniecību" [150], nākamajā desmitgadē būvniecības nozare piedzīvos ievērojamu digitālo transformāciju, kuras kulminācija būs koncepcija "Connected Construction" - pilnībā integrēta un automatizēta būvlaukuma izveide.



Attēls 9.1-13 IoT vai būvlaukuma datu ierīces var radīt un pārraidīt terabaitus datu dienā.

Digitālā būvlaukuma izveide nozīmē, ka visi būvniecības elementi - no plānošanas un loģistikas līdz darbu izpildei un kvalitātes kontrolei būvlaukumā, izmantojot fiksētās kameras un kvadrokopterus - tiks integrēti vienotā dinamiskā digitālā ekosistēmā. Iepriekš, šīs grāmatas 7. daļā, mēs jau aplūkojām Apache NiFi (7.4-5. attēls) - bezmaksas un atvērtā pirmkoda rīka iespējas, kas nodrošina reāllaika datu plūsmu - no datu vākšanas no dažādiem avotiem līdz pārsūtīšanai uz glabāšanas vai analīzes platformām.

Dati par būvniecības gaitu, materiālu patēriņu, aprīkojuma stāvokli un drošību reāllaikā tiks nosūtīti uz analītiskajām sistēmām (9.1-13. attēls). Tas ļauj prognozēt iespējamos riskus, nekavējoties reaģēt uz novirzēm un optimizēt būvlaukuma procesus. Digitālās būvlaukuma galvenās sastāvdaļas ir šādas:

- IoT - sensori - vides parametru izsekošana, celtniecības tehnikas uzraudzība un darba apstākļu kontrole.
- Digitālie dvīņi - virtuāli ēku un infrastruktūras modeļi, lai paredzētu iespējamās novirzes un novērstu kļūdas.
- Automatizētas loģistikas sistēmas - reāllaika piegādes ķēdes pārvaldība, lai samazinātu dīkstāves laiku un izmaksas.
- Robotizētas būvniecības sistēmas - autonomo mašīnu izmantošana ikdienas un bīstamu uzdevumu veikšanai.

Robotizācija, lietu interneta IoT) un digitālās būvlaukuma (būvlaukuma) koncepcijas "Connected Site (Construction)" izmantošana ne tikai palielinās efektivitāti un samazinās izmaksas, bet arī ieviešīs jaunu drošības, ilgtspējīgas būvniecības un prognozējamas projektu pārvaldības ēru.

RFID (Radio Frequency Identification) tagi ir arī viena no svarīgākajām IoT sastāvdaļām. Tās tiek izmantotas, lai identificētu un izsekotu materiālus, tehniku un pat personālu būvlaukumā, palielinot projekta resursu pārredzamību un kontroli.

RFID -tehnoloģija tiek izmantota, lai automātiski atpazītu objektus, izmantojot radio signālus. To veido trīs galvenie elementi:

- RFID - birkas (pasīvās vai aktīvās) - satur unikālu identifikatoru un tiek piestiprinātas materiāliem, instrumentiem vai iekārtām.
- Skeneri ir ierīces, kas nolasa informāciju no markējumiem un nosūta to uz sistēmu.
- Centralizēta datubāze - tajā tiek glabāta informācija par objektu atrašanās vietu, statusu un kustību.

RFID izmantošana būvniecībā:

- Automātiska materiālu uzskaite - birkas uz transportbetona izstrādājumu, armatūras vai sendvičpaneļu iepakojumiem ļauj kontrolēt krājumus un novērst zādzības.
- Personāla darba kontrole - RFID - darbinieku žetonos tiek reģistrēts maiņas sākuma un beigu laiks, nodrošinot darba laika uzskaiti.
- Iekārtu uzraudzība - RFID - sistēma seko iekārtu kustībai, novēršot dīkstāvi un uzlabojot logistikas efektivitāti.

Šo tehnoloģiju kopumu papildina uz blokķēdēm balstīti viedie līgumi, kas automatizē maksājumus, piegādes kontroli un līgumu ievērošanu bez starpniekiem, tādējādi samazinot krāpšanas un kavēšanās risku.

Šobrīd, kad nav vienota datu modeļa, viedie līgumi ir vienkārši kods, par kuru dalībnieki vienojas. Tomēr, izmantojot uz datiem vērstu pieeju, ir iespējams izveidot vienotu līguma parametru modeli, kodēt to blokķēdē un automatizēt nosacījumu izpildi.

Piemēram, piegādes ķēdes pārvaldības sistēmā viedlīgums var izsekot sūtījuma piegādei, izmantojot IoT sensorus un RFID birkas, un automātiski pārskaitīt maksājumu, kad sūtījums pienāk. Līdzīgi arī būvlaukumā viedlīgums var reģistrēt darba posma pabeigšanu, piemēram, armatūras uzstādīšanu vai pamatu ielejšanu, pamatojoties uz datiem no droniem vai būvniecības sensoriem, un automātiski iniciēt nākamo maksājumu darbuzņēmējam, neveicot manuālas pārbaudes un nepieprasot papīra sertifikātus.

Tomēr, neraugoties uz jaunajām tehnoloģijām un starptautisko standartizācijas organizāciju centieniem, lietu interneta situāciju sarežģī daudz konkurējošu standartu.

Saskaņā ar 2017. gadā publicēto Cisco pētījumu [151] gandrīz 60% lietu interneta iniciatīvu (IoT) apstājas koncepcijas pārbaudes posmā, un tikai 26% uzņēmumu uzskata, ka to IoT projekti ir pilnībā veiksmīgi. Turklat trešdaļa pabeigto projektu nesasniedz izvirzītos mērķus un netiek atzīti par veiksmīgiem pat pēc īstenošanas.

Viens no galvenajiem iemesliem ir tas, ka trūkst sadarbspējas starp platformām, kas apstrādā datus no dažādiem sensoriem. Rezultātā dati paliek izolēti atsevišķos risinājumos. Alternatīva šādai pieejai, tāpat kā citos līdzīgos gadījumos (par kuriem mēs esam rakstījuši šajā grāmatā), ir arhitektūra, kas balstīta uz pašiem datiem kā galveno aktīvu.

IoT sensoriem ir būtiska nozīme ne tikai iekārtu tehniskā stāvokļa uzraudzībā, bet arī prognozēšanas analīzē, lai samazinātu riskus būvlaukumā un uzlabotu vispārējo procesu veikspēju, paredzot klūmes un novirzes.

IoT sensoru un RFID tagu savāktos datus reāllaikā var apstrādāt ar mašīnmācīšanās algoritmiem, kas spēj noteikt anomālijas un iepriekš brīdināt inženierus par iespējamiem darbības traucējumiem. Tās var būt gan mikroplaisas betona konstrukcijās, gan netipiskas pauzes torņa celtņa darbībā, kas norāda uz tehniskām klūmēm vai normatīvo aktu pārkāpumiem. Turklāt progresīvi uzvedības analīzes algoritmi var fiksēt uzvedības modeļus, kas var liecināt, piemēram, par personāla fizisko nogurumu, tādējādi uzlabojot proaktīvu drošības un darbinieku labsajūtas pārvaldību būvlaukumā.

Būvniecības nozarē negadījumi un avārijas - gan ar tehniku, gan cilvēkiem - reti notiek pēkšni. Parasti pirms tām notiek nelielas novirzes, kas paliek nepamanītas. Prognozēšanas analītika un mašīnmācīšanās ļauj atklāt šos signālus agrīnā stadijā, vēl pirms rodas kritiskas sekas.

Būvniecības projektu digitālo nospiedumu veido dokumenti, projektu datnes un dati no IoT ierīcēm un RFID tagiem, bet mašīnmācīšanās var palīdzēt iegūt no tiem noderīgu informāciju. Pieaugot datu apjomam un demokratizējoties datu piekļuvei, būvniecības nozarei paveras jaunas iespējas analītikas, prognozēšanas analītikas un mākslīgā intelekta lietojumu jomā.

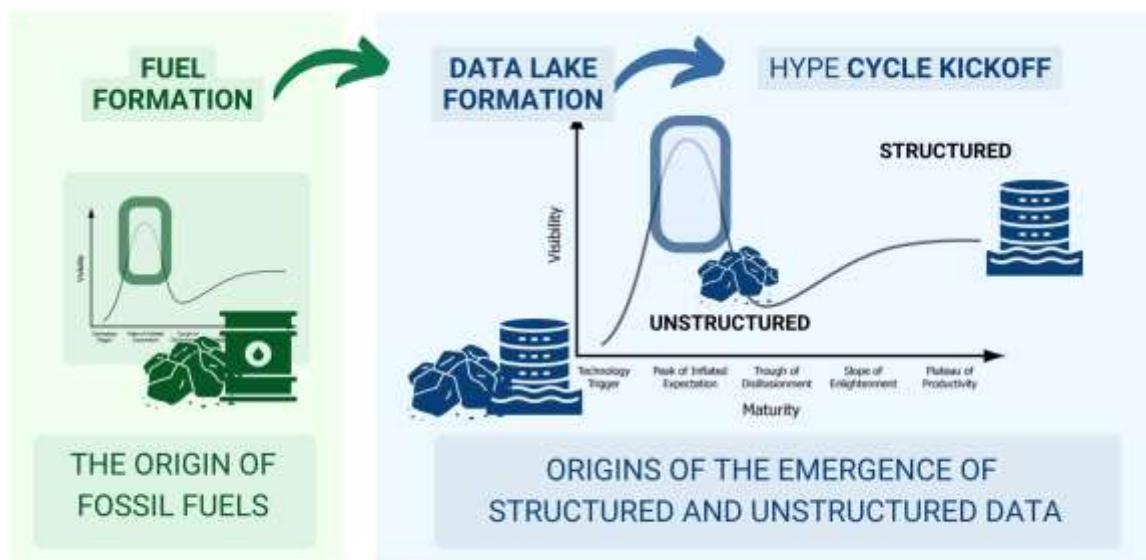


## NODAĻA 9.2.

### MAŠĪNMĀCĪŠANĀS UN PROGNOZES

#### Mašīnmācīšanās un mākslīgais intelekts mainīs veidu, kā mēs veidojam.

Dažādu būvniecības nozares sistēmu datubāzes ar to neizbēgami bojājošo un arvien sarežģītāko infrastruktūru kļūst par augsti nākotnes risinājumu izstrādei. Uzņēmumu serveri, līdzīgi kā mežs, ir bagāti ar svarīgas informācijas biomasu, kas bieži vien paslēpta zem zemes, mapju un serveru dzīlēs. Datu masas no dažādām sistēmām, kas tiek radītas šodien pēc lietošanas, pēc nokrišanas servera apakšā un pēc gadiem ilgas fosilizācijas - nākotnē kalpos kā degviela mašīnmācīšanās un valodas modeļiem. Uzņēmumu iekšējās tērzētavas (piemēram, atsevišķa lokāli konfigurēta ChatGPT, LlaMa, Mistral, DeepSeek instance) tiks veidotas, izmantojot šos iekšējos modeļus, izmantojot centralizētu krātuvi, lai ātri un ērti izgūtu informāciju un ģenerētu nepieciešamos grafikus, paneļus un dokumentus.



Attēls 9.2-1 Tāpat kā koki laika un analītisko datu ietekmē laika gaitā pārvēršas oglēs, tā arī informācija laika un analītisko datu ietekmē laika gaitā pārvēršas vērtīgā biznesa energijā.

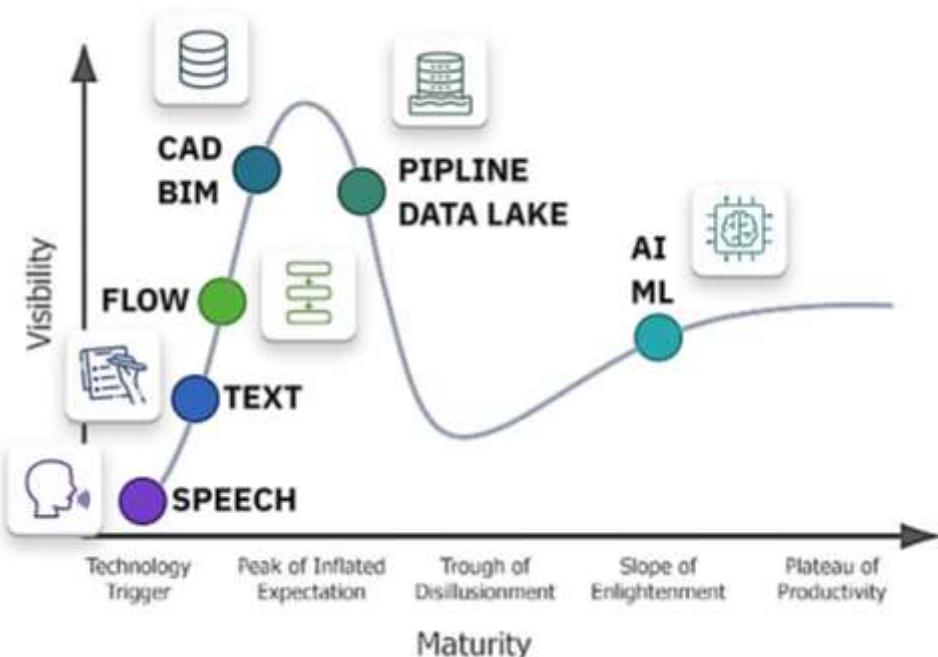
Augu masas fosilizācija spiediena un temperatūras ietekmē rada viendabīgu un unikālu strukturētu viendabīgu dažādu sugu koku, kas dzīvojuši dažādos laikos, masu - kokogles [152]. Līdzīgā veidā informācija, kas dažādos formātos un dažādos laikos ierakstīta cietajos diskos analītisko nodaļu spiediena un kvalitātes vadības temperatūras ietekmē, galu galā veido viendabīgu strukturētu vērtīgas informācijas homogēnu masu (9.2-1. attēls).

Šos informācijas slāņus (vai biežāk izolētus informācijas gabaliņus) rada pieredzējuši analītiķi, rūpīgi organizējot datus, un viņi pakāpeniski sāk izvilkst vērtīgu informāciju no šķietami ilgi nebūtiskiem datiem.

Brīdī, kad šie nobriedušie datu slāņi vairs netiek tikai "sadedzināti" pārskatos, bet sāk cirkulēt biznesa procesos, bagātinot lēmumus un uzlabojot procesus, uzņēmums kļūst gatavs nākamajam solim -

pārejai uz mašīnmācīšanos un mākslīgo intelektu ( 9.2-2. att.).

Mašīnmācīšanās (ML - Machine learning) ir mākslīgā intelekta problēmu risināšanas metožu klase. Mašīnmācīšanās algoritmi atpazīst modeļus lielās datu kopās un izmanto tos, lai paši mācītos. Katra jauna datu kopa ļauj matemātiskajiem algoritmiem pilnveidoties un pielāgoties atbilstoši iegūtajai informācijai, kas ļauj pastāvīgi uzlabot ieteikumu un prognožu precizitāti.



Attēlā: 9.2-2. attēls Datu radīšanas tehnoloģiju izsušana un analītisko rīku izmantošana paver durvis mašīnmācīšanās tēmai.

Kā 2023. gada intervijā teica ietekmīgā pasaules lielākā ieguldījumu fonda (kam pieder galvenās daļas gandrīz visos lielākajos būvniecības programmatūras uzņēmumos, kā arī uzņēmumos, kuriem pieder lielākais nekustamā īpašuma apjoms pasaulē [55]) vadītājs. - mašīnmācīšanās mainīs būvniecības pasauli.

*AI ir milzīgs potenciāls. Tas mainīs mūsu darba un dzīvesveidu. Mākslīgais intelekts un robotika mainīs veidu, kā mēs strādājam un būvējam, un mēs varēsim izmantot mākslīgo intelektu un robotiku kā līdzekli, lai radītu daudz lielāku produktivitāti [153].*

- Pasaulē lielākā ieguldījumu fonda izpilddirektors, intervija, 2023. gada septembris.

Mašīnmācīšanās (ML) darbojas, apstrādājot lielus datu apjomus, izmantojot statistikas metodes, lai atdarinātu cilvēka domāšanas aspektus. Tomēr lielākajai daļai uzņēmumu nav šādu datu kopu, un, ja ir, tad tās bieži vien nav pietiekami markētas. Šeit var palīdzēt semantiskās tehnoloģijas un pārneses

mācīšanās - tehnika, kas ļauj ML efektīvāk apstrādāt nelielus datu apjomus, kuras iespējamība ir aplūkota šīs daļas iepriekšējās nodaļās.

Pārneses mācīšanās būtība ir tāda, ka tā vietā, lai mācītos katru uzdevumu no jauna, jūs varat izmantot zināšanas, kas iegūtas saistītās jomās. Ir jāapzinās, ka citās nozarēs iegūtos modeļus un atklājumus var pielāgot un pielietot būvniecības nozarē. Piemēram, mazumtirdzniecībā izstrādātās loģistikas procesu optimizācijas metodes palīdz uzlabot būvniecības piegādes kēdes pārvaldības efektivitāti. Lielo datu analīzi, ko aktīvi izmanto finansēs, var piemērot izmaksu prognozēšanai un risku pārvaldībai būvniecības projektos. Savukārt datorredzes un robotikas tehnoloģijas, kas tiek attīstītas rūpniecībā, jau atrod pielietojumu automatizētā kvalitātes kontrole, drošības uzraudzībā un būvlaukumu apsaimniekošanā.

Transfermācīšanās ļauj ne tikai paātrināt inovāciju ieviešanu, bet arī samazināt to izstrādes izmaksas, izmantojot jau uzkrāto citu nozaru pieredzi.

$$\text{labor productivity in construction} = f(\text{AI})$$

**Attēlā: 9.2-3 Mākslīgais intelekts tehnoloģijas un robotika būs galvenais virzītājspēks nākotnē, lai palielinātu produktivitāti būvniecības nozarē.**

Cilvēka domāšana ir organizēta pēc līdzīga principa: mēs balstāmies uz iepriekš iegūtām zināšanām, lai risinātu jaunas problēmas ( 4.4-19. attēls, 4.4-20. attēls, 4.4-21. attēls). Arī mašīnmācībā šī pieeja darbojas - vienkāršojot datu modeli un padarot to elegantāku, mēs varam samazināt problēmas sarežģītību ML algoritmiem. Tas savukārt samazina vajadzību pēc liela datu apjoma un samazina skaitlošanas izmaksas.

### No subjektīva novērtējuma līdz statistiskai prognozei

Laikmets, kad stratēģiskie lēmumi bija atkarīgi no individuālu vadītāju intuīcijas ( 9.2-4. attēls), ir pagātne. Aizvien pieaugošas konkurences un izaicinājumu pilnajā ekonomiskajā vidē subjektīva pieeja klūst pārāk riskanta un neefektīva. Uzņēmumi, kas turpina paļauties uz personīgo viedokli, nevis objektīvi analizē datus, zaudē spēju ātri reaģēt uz pārmaiņām.

Konkurences vide prasa precizitāti un atkārtojamību, pamatojoties uz datiem, statistikas modeļiem un aprēķināmu varbūtību. Lēmumi vairs nevar balstīties uz izjūtām, tiem jābalstās uz korelācijām, tendencēm un prognozēšanas modeļiem, kas iegūti, izmantojot analītiku un mašīnmācīšanos. Tā nav tikai instrumentu maiņa - tā ir domāšanas loģikas maiņa: no pieņēmumiem uz pierādījumiem, no subjektīvām varbūtībām uz statistiski aprēķinātām novirzēm, no sajūtām uz faktiem.



**Attēlā: 9.2-4 Līdz ar lielo datu un mašīnmācīšanās tehnoloģiju ienākšanu ēra, kad lēmumus pieņēma HiPPO (vislabāk atalgoņā darbinieka viedoklis), kļūs pagātne.**

Vadītāji, kuri agrāk paļāvās tikai uz savām izjūtām, neizbēgami saskarsies ar jaunu realitāti: autoritāte vairs nenosaka izvēli. Tagad vadības centrā ir sistēmas, kas analizē miljoniem parametru un vektoru, identificē slēptos modeļus un iesaka optimālas stratēģijas.

Galvenais iemesls, kāpēc uzņēmumi mūsdienās joprojām izvairās no ML ieviešanas, ir tās pārredzamības trūkums. Lielākā daļa modeļu vadītājiem darbojas kā "melnās kastes", nepaskaidrojot, kā tieši tie nonāk pie saviem secinājumiem. Tas rada problēmas: algoritmi var pastiprināt stereotipus un pat radīt humoristiskas situācijas, kā Microsoft tērzēšanas robota gadījumā, kas ātri vien pārvērtās par toksisku komunikācijas rīku [154].

Bijušais pasaules šaha čempions Garijs Kasparovs grāmatā "Dzīļā domāšana" (Deep Thinking) atspoguļo savu sakāvi ar IBM datoru Big Blue [155]. Viņš apgalvo, ka mākslīgā intelekta patiesā vērtība nav cilvēka intelekta kopēšana, bet gan mūsu spēju papildināšana. Mākslīgajam intelektam būtu jāveic uzdevumi, kuros cilvēks ir vājš, bet cilvēks sniedz radošumu. Datori ir mainījuši tradicionālo pieeju šaha analīzes veikšanai. Tā vietā, lai radītu aizraujošus stāstus par spēlēm, datoršaha programmas katru gājienu vērtē objektīvi, pamatojoties tikai uz tā faktisko spēku vai vājumu. Kasparovs norāda, ka cilvēka tieksme skatīt notikumus kā saskaņotus stāstus, nevis kā atsevišķas darbības, bieži novēd pie nepareiziem secinājumiem - ne tikai šahā, bet arī dzīvē kopumā.

Tāpēc, ja plānojat izmantot mašīnmācīšanos prognozēšanai un analīzei, ir svarīgi izprast tās pamatprincipus - kā darbojas algoritmi un kā tiek apstrādāti dati - pirms sākat savā darbā izmantot mašīnmācīšanās rīkus un mākslīgo intelektu. Vislabākais veids, kā sākt strādāt, ir gūt praktisku pieredzi.

Viens no ērtākajiem rīkiem sākotnējam ievadam mašīnmācīšanās tēmā un prognozēšanā ir Jupyter Notebook un populārais klasiskais Titanic datu kopums, kas sniegs vizuālu ievadu galvenajās datu analīzes un ML modeļu veidošanas metodēs.

## **Titānika datu kopa: Hello World analītikas pasaulē dati un lielie dati**

Viens no slavenākajiem ML izmantošanas datu analīzē piemēriem ir "Titānika" datu kopas analīze, ko bieži izmanto, lai pētītu pasažieru izdzīvošanas varbūtību. Šīs tabulas apguve ir analoga programmai "Hello World", apgūstot programmēšanas valodas.

RMS "Titāniks" nogrimšana 1912. gadā prasīja 1502 no 2224 cilvēkiem dzīvību. Titanic datu kopa satur ne tikai informāciju par to, vai pasažieris izdzīvoja, bet arī tādus raksturlielumus kā vecums, dzimums, biješu klase un citi parametri. Šīs datu kopums ir pieejams bez maksas, un to var atvērt un analizēt dažādās bezsaistes un tiešsaistes platformās.

Saitē uz Titanic datu kopu:

<https://raw.githubusercontent.com/datasets/master/titanic.csv>

Iepriekš nodaļā "IDE ar LLM atbalstu un turpmākās programmēšanas izmaiņas" mēs jau aplūkojām Jupyter Notebook - vienu no populārākajām datu analīzes un mašīnmācīšanās izstrādes vidēm. Jupyter Notebook bezmaksas mākoņa analogi ir Kaggle un Google Collab platformas, kas ļauj palaist Python kodu bez programmatūras instalēšanas un nodrošina bezmaksas piekļuvi skaitļošanas resursiem.

Kaggle ir lielākā datu analīzes un mašīnmācīšanās sacensību platforma ar integrētu koda izpildes vidi. Pēc 2023. gada oktobra datiem Kaggle ir vairāk nekā 15 miljoni lietotāju [156] no 194 valstīm.

Lejupielādējiet un izmantojet Titanic datu kopu Kaggle platformā ( 9.2-5. attēls), lai saglabātu datu

kopu (tās kopiju) un palaistu Python kodu ar iepriekš instalētām bibliotēkām tieši pārlūkprogrammā, bez nepieciešamības instalēt ūpašu IDE.



Attēls 9.2-5 Titanic tabulas statistika - populārākā mācību datu kopa datu analīzes un mašīnmācīšanās apguvei.

"Titānika" datu kopa ietver datus par 2224 pasažieriem, kas atradās uz RMS "Titāniks" tā bojāejas brīdī 1912. gadā. Datu kopa ir iesniegta kā divas atsevišķas tabulas - mācību (train.csv) un testa (test.csv) paraugs, kas ļauj to izmantot gan modeļu apmācībai, gan to precizitātes novērtēšanai, izmantojot jaunus datus.

Mācību datu kopa satur gan pasažieru atribūtus (vecums, dzimums, biļešu klase un citi), gan informāciju par izdzīvojušajiem (sleja ar bināro vērtību "Izdzīvoja"). Mācību datu kopu (9.2-6. attēls - fails train.csv) izmanto, lai apmācītu modeli. Testa datu kopa (9.2-7. attēls - fails test.csv) ietver tikai pasažieru atribūtus bez informācijas par izdzīvojušajiem (bez vienas ailes "Izdzīvojušais"). Testa datu kopa ir paredzēta, lai pārbaudītu modeli ar jauniem datiem un novērtētu tā precizitāti.

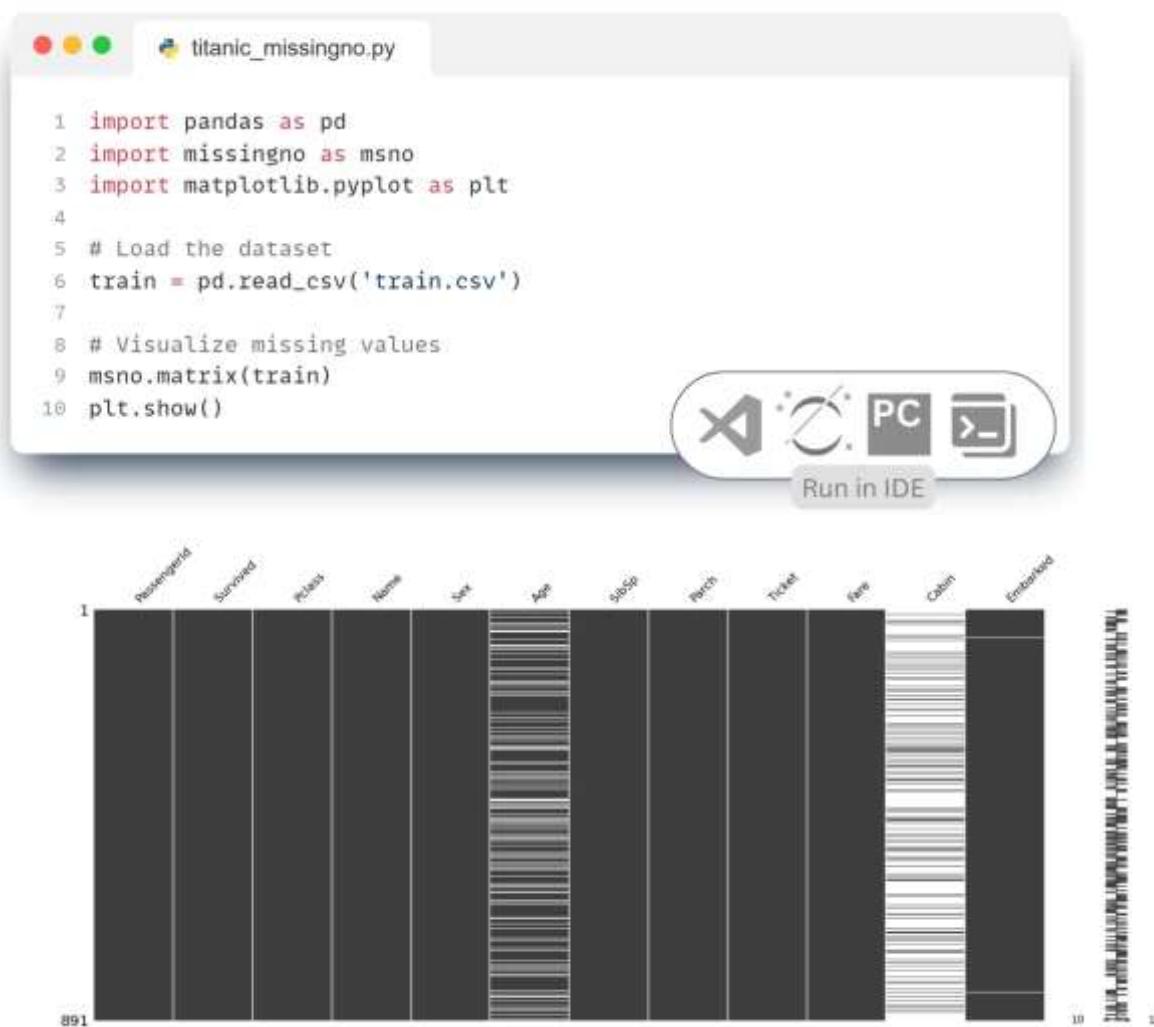
Tādējādi mums ir gandrīz identiski pasažieru atribūti mācību un testa datu kopās. Vienīgā būtiskākā atšķirība ir tā, ka testa datu kopā mums ir saraksts ar pasažieriem, kuriem nav slejas "Izdzīvojušais" - mērķa mainīgais, kuru mēs vēlamies iemācīties prognozēt, izmantojot dažādus matemātiskus algoritmus. Un pēc modeļa izveides mēs varēsim salīdzināt mūsu modeļa rezultātus ar reālo parametru "Survivor" no testa datu kopas, ko ņemsim vērā, lai novērtētu rezultātus.

Tabulas galvenās kolonas, pasažieru parametri mācību un testa datu kopā:

- **PassengerId** - unikāls pasažiera identifikatorς
- **Izdīvojis** - 1, ja pasažieris izdzīvojis, 0, ja miris (nav pieejams testa kopā).
- **Pclass** - biļešu klase (1, 2 vai 3).
- **Vārds** - pasažiera vārds un uzvārds

- **Dzimums** - pasažiera dzimums (vīrietis/sieviete).
- Vecums
- **SibSp** - brāļu/māsu vai laulāto skaits uz kuģa
- **Parch** - vecāku vai bērnu skaits uz kuģa
- **Biljetē** - biljetes numurs
- **Tarifs** - biljetes cena
- Kabīne - kabīnes numurs (trūkst daudzu datu)
- **Iekāpšana** ir iekāpšanas osta (C = Cherbourg, Q = Queenstown, S = Southampton).

Lai vizualizētu trūkstošos datus abās tabulās, varat izmantot bibliotēku missingno (9.2-6. attēls, 9.2-7. attēls), kas attēlo trūkstošās vērtības histogrammas veidā, kur baltie lauki parāda trūkstošos datus. Šī vizualizācija ļauj ātri novērtēt datu kvalitāti pirms apstrādes.



**Attēls 9.2-6** Dažas koda rindas tiek izmantotas, lai vizualizētu trūkstošos datus "Titānika" mācību datu kopā, kur galvenais parametrs apmācībai ir parametrs "Izdzīvoti".



**Attēls 9.2-7** Vizualizācija trūkstošo datu testa datu kopā Titanic, kurā ir tikai pasažieru raksturlielumi bez informācijas.

Pirms hipotēžu formulēšanas un prognožu veikšanas, pamatojoties uz datu kopu, vizuālā analīze palīdz identificēt galvenos datu modeļus, novērtēt to kvalitāti un noteikt iespējamās atkarības. Ir daudz vizualizācijas metožu, kas var palīdzēt labāk izprast "Titānika" datu kopu. Jūs varat izmantot sadalījuma diagrammas, lai analizētu pasažieru vecuma grupas, izdzīvošanas diagrammas pēc dzimuma un klases, kā arī trūkstošo datu matricas, lai novērtētu informācijas kvalitāti un izprastu datus.

- Palūgsim LLM palīdzēt mums vizualizēt datus no Titanic datu kopas, nosūtot šādu teksta pieprasījumu jebkuram LLM modelim (CHATGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN vai jebkuram citam):

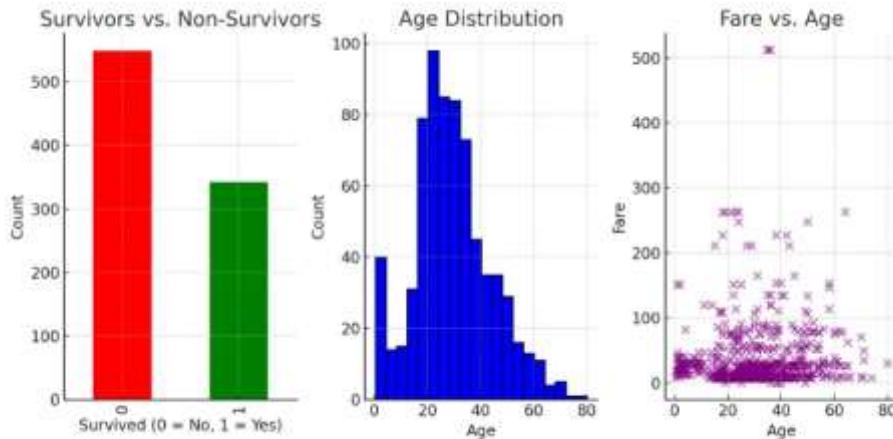
Lūdzu, parādiet dažus vienkāršus grafikus par Titānika datu kopu. Paši lejupielādējiet datu kopu un parādiet ↴

- LLM atbilde gatava koda un grafiku veidā, kas vizualizē datu kopas parametrus

```

1 import pandas as pd
2 import matplotlib.pyplot as plt
3
4 # Load the dataset
5 file_path = "/mnt/data/titanic.csv"
6 df = pd.read_csv(file_path)
7
8 # Convert column names to lowercase for consistency
9 df.columns = df.columns.str.lower()
10
11 # Basic Plots
12 plt.figure(figsize=(10, 5))
13
14 # Bar Chart: Survival Count
15 plt.subplot(1, 3, 1)
16 df['survived'].value_counts().plot(kind='bar', color=['red', 'green'])
17 plt.title("Survivors vs. Non-Survivors")
18 plt.xlabel("Survived (0 = No, 1 = Yes)")
19 plt.ylabel("Count")
20
21 # Histogram: Age Distribution
22 plt.subplot(1, 3, 2)
23 df['age'].dropna().plot(kind='hist', bins=20, color='blue', edgecolor='black')
24 plt.title("Age Distribution")
25 plt.xlabel("Age")
26 plt.ylabel("Count")
27
28 # Scatter Plot: Fare vs. Age
29 plt.subplot(1, 3, 3)
30 plt.scatter(df['age'], df['fare'], alpha=0.5, color='purple')
31 plt.title("Fare vs. Age")
32 plt.xlabel("Age")
33 plt.ylabel("Fare")
34
35 # Show the plots
36 plt.tight_layout()
37 plt.show()

```



Attēls 9.2-8 LLM palīdz iegūt tūlītēju datu kopas datu vizualizāciju.

Datu vizualizācija ir svarīgs solis, lai sagatavotu datu kopu turpmākai mašīnmācīšanās modeļa izveidei, ko var veikt, tikai izprotot datus.

## Mašīnmācīšanās darbībā: no "Titānika" pasažieriem līdz projektu pārvaldībai

Galvenā hipotēze, kas izmantota, lai izpētītu mašīnmācīšanās sistēmu, pamatojoties uz "Titānika" datu kopu, ir tāda, ka noteiktām pasažieru grupām bija lielākas izdzīvošanas izredzes.

Titānika pasažieru mazā tabula ir kļuvusi populāra visā pasaulei, un miljoniem cilvēku to izmanto apmācībai, eksperimentiem un modeļu testēšanai, lai noskaidrotu, kuri algoritmi un hipotēzes izveidos visprecīzāko izdzīvošanas prognozēšanas modeli, pamatojoties uz Titānika pasažieru mācību datu kopu.

Titānika datu kopas pievilcība ir saistīta ar tās kompaktumu: ar vairākiem simtiem rindu un divpadsmit kolonnām (9.2-6. attēls) tā sniedz plašas analīzes iespējas. Datu kopa salīdzinoši vienkārši ir klasisks bināro klasifikācijas risinājumu piemērs, kur problēmas mērķis - izdzīvošana - ir izteikts ērtā formātā 0 vai 1.

Džons Vīlers filmā "It from Bit" [7] apgalvo, ka Visuma pamatā ir bināra izvēle. Līdzīgi arī bizness, ko vada cilvēki, kas sastāv no molekulām, patiesībā ir balstīts uz virkni bināro bināro izvēju.

Turklāt dati ir balstīti uz reālu vēsturisku notikumu, kas atšķirībā no mākslīgi radītiem piemēriem tos padara vērtīgus pētniecībai. Tikai Kaggle platformā, kas ir viena no lielākajām datu cauruļvadu un ETL, uz Titānika datu kopu balstītos uzdevumos piedalījās 1 355 998 cilvēki, izstrādājot 53 963 unikālus datu cauruļvadu risinājumus [157] (9.2-9. attēls).

Šķiet neticami, bet tikai 1000 rindiņu datu par Titānika pasažieriem ar 12 parametriem ir kļuvuši par lauku miljoniem hipotēžu, logisko kēžu un unikālu datu cauruļvadu. No nelielas datu kopas dzimst bezgalīgas atziņas, hipotēzes un interpretācijas - no vienkāršiem izdzīvošanas modeļiem līdz sarežģītiem ansambljiem, kuros ņemti vērā slēptie modeļi un sarežģīti domāšanas labirinti.

**Machine Learning from Disaster**

[Submit Prediction](#)

Data    Code    Models    Discussion    Leaderboard    Rules

---

**Titanic Tutorial**    16916  
 Updated 3y ago    Gold    ---  
 29858 comments · Titanic - Machine Learning from Disaster

---

**Titanic competition w/ TensorFlow Decision Forests**    1098  
 Updated 2y ago    Gold    ---  
 Score: 0.80143 · 318 comments · Titanic - Machine Learning from Disaster

---

**Titanic Data Science Solutions**    10723  
 Updated 6y ago    Gold    ---  
 2590 comments · Titanic - Machine Learning from Disaster

---

**Exploring Survival on the Titanic**    3968  
 Updated 7y ago    Gold    ---  
 Score: 0.80382 · 1072 comments · Titanic - Machine Learning from Disaster

Attēls 9.2-9 Pirmie pieci risinājumi no kopumā 53 963 gataviem un atvērtā pirmkoda risinājumiem Pipeline. Šo problēmu Kaggle vietnē vien jau ir mēģinājuši atrisināt gandrīz 1,5 miljoni cilvēku [157].

Ja pat šāda neliela tabula var ģenerēt miljoniem unikālu risinājumu ( 9.2-9. attēls), ko mēs varam teikt par reālām rūpnieciskās būvniecības datu kopām, kurās parametri mērīti desmitos tūkstošu?

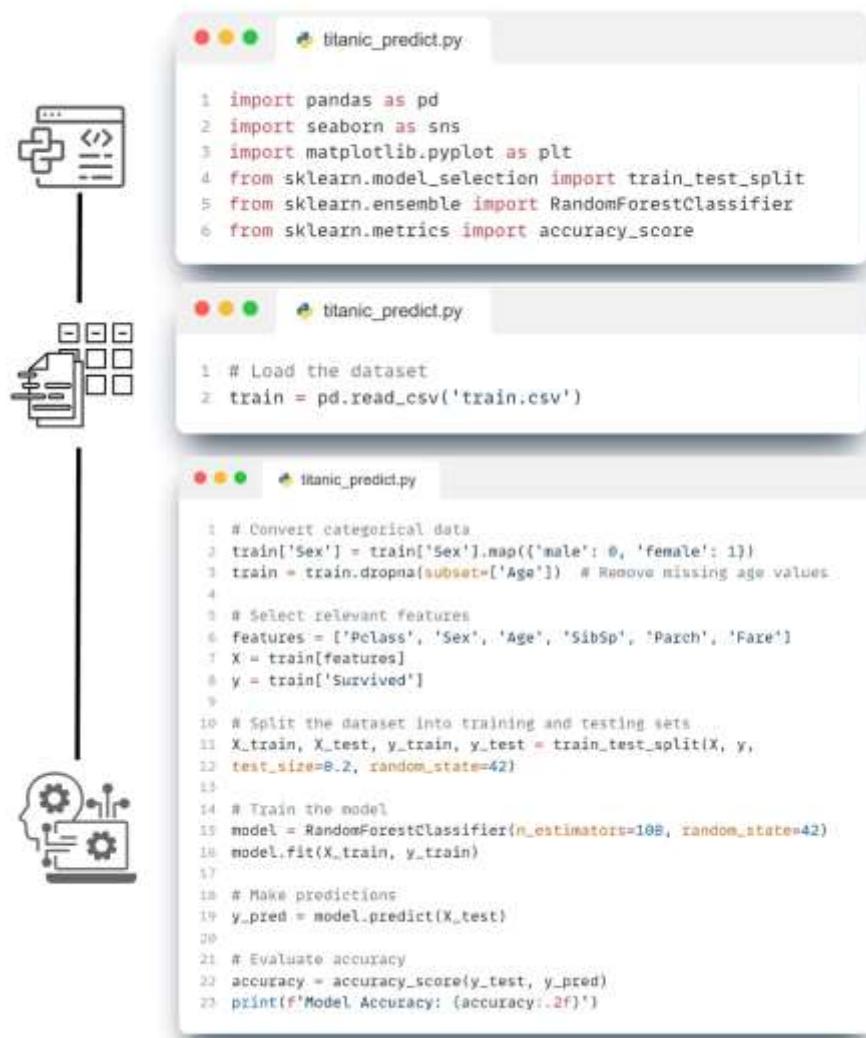
Standarta CAD - salīdzinoši nelielas ēkas projektā ir desmitiem tūkstošu vienību ar tūkstošiem parametru - no ģeometriskiem parametriem līdz izmaksu un laika atribūtiem. Iedomājieties, cik daudz potenciālu atziņu, sakarību, prognožu un vadības hipotēžu slēpjās visu jūsu uzņēmuma gadu gaitā savākto projektu datos. Vēsturiskie projektu dati nav tikai arhīvs - tā ir organizācijas dzīvā atmiņa, tās digitālais nos piedums, ko var analizēt, lai radītu lielu skaitu unikālu hipotēžu.

Vissvarīgākais ir tas, ka jums nav jāgaida, kad Kaggle kopiena izrādīs interesu par jūsu uzņēmumu vai jūsu datiem. Jūs varat sākt strādāt ar to, kas jums ir jau šodien: palaist analīzi saviem datiem, apmācīt modeļus saviem datiem, noteikt atkārtojumus, anomālijas un modeļus. Ja agrāk bija nepieciešami gadiem ilgi eksperimenti un dārgas konsultācijas, tagad pietiek ar iniciatīvu, LLM, atvērtu pieejumu datiem un vēlmi mācīties.

- ❷ Lai izveidotu mašīnmācīšanās algoritmu, kas prognozēs pasažieru izdzīvošanas rādītājus, pamatojoties uz train.csv pasažieru mācību datu kopu, lūgsim LLM atrisināt šo problēmu mūsu vietā:

Pamatojoties uz "Titānika" pasažieru mācību datu kopu, izveidojiet mašīnmācīšanās modeli, lai prognozētu izdzīvošanas rādītājus. ↵

 LLM atbilde :



```

1  import pandas as pd
2  import seaborn as sns
3  import matplotlib.pyplot as plt
4  from sklearn.model_selection import train_test_split
5  from sklearn.ensemble import RandomForestClassifier
6  from sklearn.metrics import accuracy_score

1 # Load the dataset
2 train = pd.read_csv('train.csv')

1 # Convert categorical data
2 train['Sex'] = train['Sex'].map({'male': 0, 'female': 1})
3 train = train.dropna(subset=['Age']) # Remove missing age values
4

5 # Select relevant features
6 features = ['Pclass', 'Sex', 'Age', 'SibSp', 'Parch', 'Fare']
7 X = train[features]
8 y = train['Survived']
9

10 # Split the dataset into training and testing sets ...
11 X_train, X_test, y_train, y_test = train_test_split(X, y,
12 test_size=0.2, random_state=42)
13

14 # Train the model
15 model = RandomForestClassifier(n_estimators=100, random_state=42)
16 model.fit(X_train, y_train)
17

18 # Make predictions
19 y_pred = model.predict(X_test)
20

21 # Evaluate accuracy
22 accuracy = accuracy_score(y_test, y_pred)
23 print(f'Model Accuracy: {accuracy:.2f}')

```

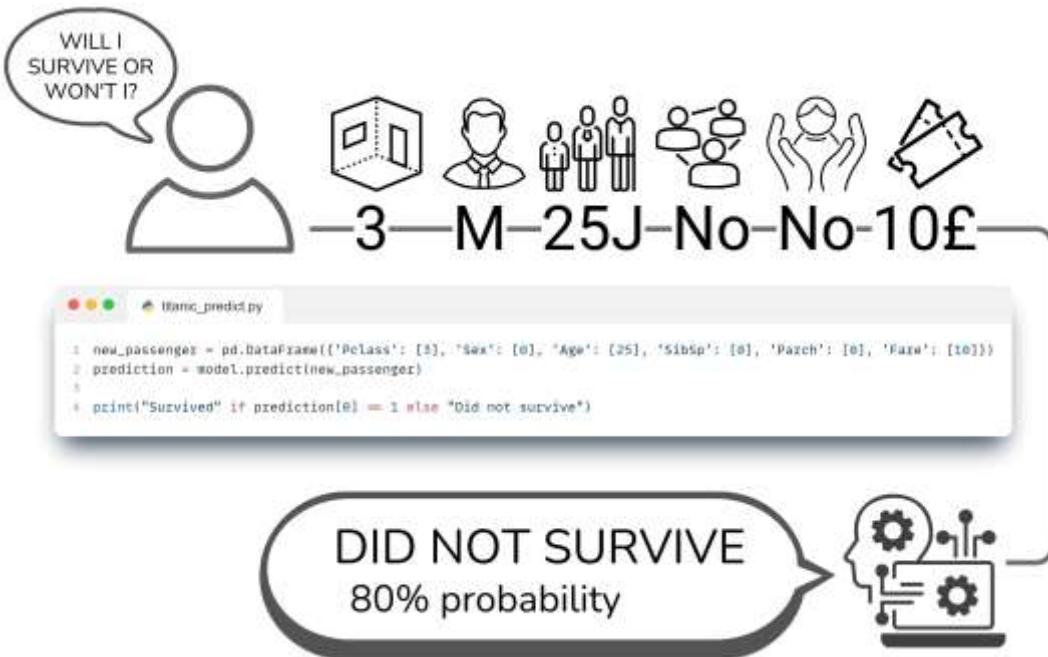
Attēlā 9.2-10 LLM izveidoja prognozēšanu par Titānika izdzīvojušajiem, izmantojot mašīnmācīšanās algoritmu Random Forest.

legūtais LLM kods (9.2-10. attēls) ielādē Titānika pasažieru datus, attīra tos, konvertē kategoriskos mainīgos (piemēram, dzimumu skaitliskajā formātā) un trenē modeli, izmantojot RandomForestClassifier algoritmu, lai prognozētu, vai pasažieris izdzīvoja vai nē (par populārākajiem algoritmiem sīkāk runāsim nākamajās nodaļās).

Kods sadala apmācības datus apmācības un testa kopās (Kaggle vietnē jau ir izveidoti gatavi test.csv (9.2-7. attēls) un train.csv (9.2-6. attēls) apmācībai, pēc tam modelis tiek apmācīts uz apmācības datiem un testēts uz testa datiem, lai redzētu, cik labs ir konkrētais prognozēšanas modelis. Pēc apmācības modelim tiek ievadīti testa dati no test.csv (ar reāliem datiem par izdzīvojušajiem un neizdzīvojušajiem), un tas prognozē, kurš izdzīvoja un kurš ne. Mūsu gadījumā iegūtā mašīnmācīšanās modeļa precīzitāte ir aptuveni 80%, kas liecina, ka tas diezgan labi uztver modeļus.

Mašīnmācīšanos var salīdzināt ar bērnu, kas mēģina ievietot taisnstūra klucīti apaļā caurumā. Sākotnējos posmos algoritms izmēģina daudzas pieejas, sastopoties ar kļūdām un neatbilstībām. Šis process var šķist neefektīvs, taču tas nodrošina svarīgu mācīšanos: analizējot katru kļūdu, modelis uzlabo savus paredzējumus un pieņem arvien precīzākus lēmumus.

Tagad šo modeli (9.2-10. attēls) var izmantot, lai prognozētu jaunu pasažieru izdzīvošanas rādītājus, un, piemēram, ja tam tiek ievadīta informācija par pasažieri, izmantojot funkciju model.predict ar parametriem: "vīrietis", "3. klase", "25 gadi", "nav radinieku uz kuģa", modelis sniegs prognozi - ka pasažieris ar 80% varbūtību neizdzīvos katastrofu, ja viņš atradīsies uz "Titānika" 1912. gadā (9.2-11. att.).



Attēls 9.2-11 leprieķš izveidotais modelis tagad ar 80% varbūtību var prognozēt, vai kāds no jaunajiem "Titānika" pasažieriem izdzīvos.

Titānika pasažieru izdzīvošanas prognozēšanas modelis ilustrē daudz plašāku jēdzienu: katru dienu tūkstošiem būvniecības nozares profesionāļu pieņem līdzīgus "divējādus" lēmumus - lēmuma, projekta, tāmes, rīka, peļņas vai zaudējumu, drošības vai riska, dzīvības vai nāves. Tāpat kā "Titānika" piemērā, kur iznākums bija atkarīgs no faktoriem (dzimums, vecums, klase), arī būvniecībā katru

lēmuma aspektu ietekmē daudzi savi faktori un mainīgie (tabulu kolonas): materiālu izmaksas, strādnieku kvalifikācija, termiņi, laikapstākļi, logistika, tehniskie riski, komentāri un simtiem tūkstošu citu parametru.

Būvniecības nozarē mašīnmācīšanās darbojas pēc tādiem pašiem principiem kā citās jomās: modeļi tiek apmācīti, izmantojot vēsturiskos datus - no projektiem, līgumiem, tāmēm -, lai pārbaudītu dažādas hipotēzes un atrastu efektīvākos risinājumus. Šis process ir līdzīgs bērna mācīšanai, izmantojot izmēģinājumus un klūdas: ar katru ciklu modeļi pielāgojas un klūst precīzāki.

Uzkrāto datu izmantošana paver jaunus apvāršņus būvniecībā. Tā vietā, lai veiktu laikietlpīgus manuālus aprēķinus, var apmācīt modeļus, kas spēj ar augstu precīzitāti prognozēt nākotnes projektu galvenos parametrus. Šādā veidā prognozējošā analītika pārveido būvniecības nozari par telpu, kurā var ne tikai plānot, bet arī droši prognozēt attīstību.

## **Prognozes un prognozes, pamatojoties uz vēsturiskiem datiem**

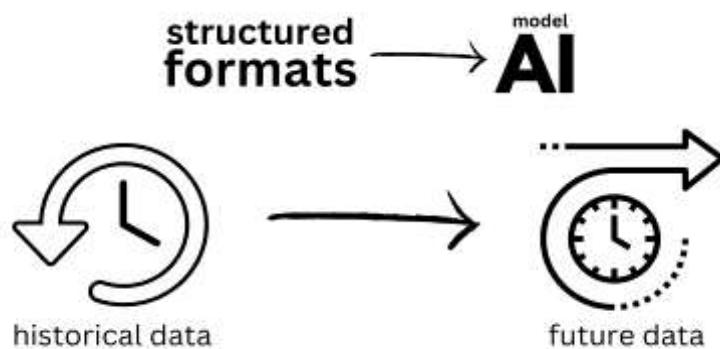
Uzņēmuma projektos savāktie dati paver iespēju izveidot modeļus, kas ļauj prognozēt nākotnes, vēl nerealizētu objektu izmaksu un laika rādītājus - bez laikietlpīgiem manuāliem aprēķiniem un salīdzinājumiem. Tas ļauj ievērojami paātrināt un vienkāršot vērtēšanas procesus, paļaujoties nevis uz subjektīviem pieņēmumiem, bet uz pamatotām matemātiskām prognozēm.

Iepriekš, grāmatas ceturtajā daļā, mēs detalizēti aplūkojām tradicionālās projektu izmaksu aplēses metodes, tostarp uz resursiem balstīto metodi, kā arī minējām parametriskās un ekspertu pieejas. Šīs metodes joprojām ir aktuālas, taču mūsdienu praksē tās sāk bagātināt ar statistiskās analīzes un mašīnmācīšanās rīkiem, kas var ievērojami uzlabot aplēšu precīzitāti un atkārtojamību.

Manuālo un pusautomātisko cenu un laika atribūtu aprēķināšanas procesu nākotnē papildinās ML modeļu viedoklis un prognozes, kas spēj analizēt vēsturiskos datus, atrast slēptos modeļus un ierosināt pamatotus lēmumus.

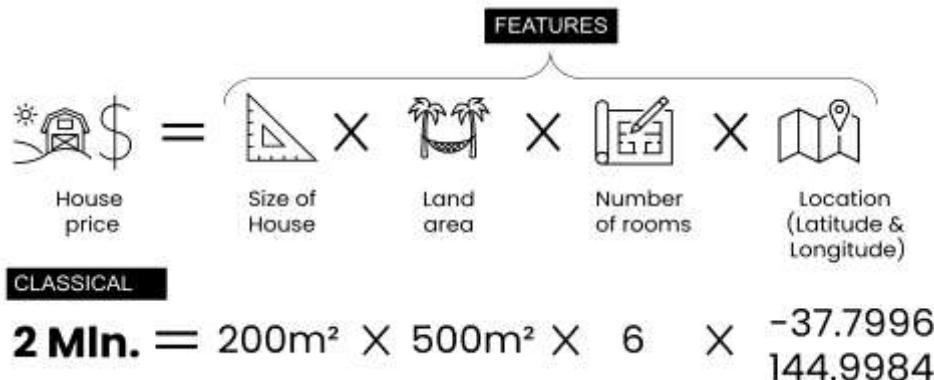
Jauni dati un scenāriji tiks automātiski generēti no jau pieejamās informācijas - līdzīgi kā valodas modeļi (LLM) veido tekstus, attēlus un kodus, pamatojoties uz datiem, kas gadiem ilgi vākti no atvērtiem avotiem [158].

Tāpat kā šodien cilvēki, lai novērtētu nākotnes notikumus, paļaujas uz pieredzi, intuīciju un iekšējo statistiku, tuvākajos gados būvniecības projektu nākotni arvāk noteiks uzkrāto zināšanu un matemātisko mašīnmācīšanās modeļu kombinācija.



**Attēls 9.2-12 Kvalitatīvi un strukturēti uzņēmuma vēsturiskie dati ir materiāls, uz kura pamata tiek veidoti mašīnmācīšanās modeļi un prognozes.**

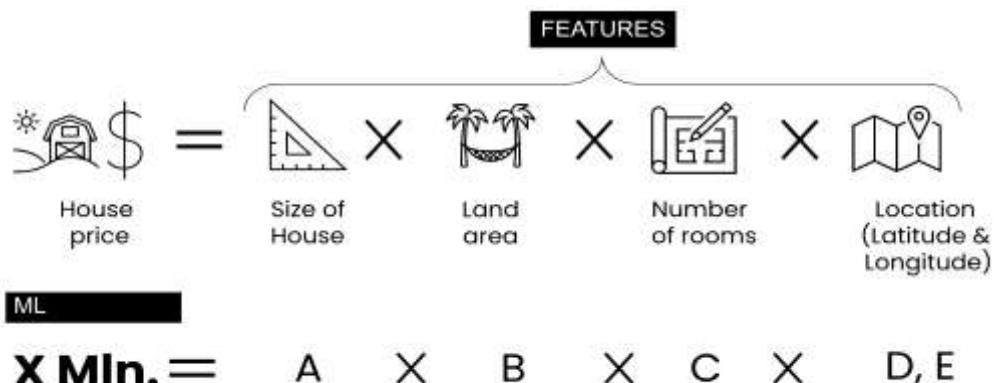
Aplūkosim vienkāršu piemēru: mājas cenas prognozēšana, pamatojoties uz tās platību, zemes gabala lielumu, istabu skaitu un ģeogrāfisko atrašanās vietu. Viena no pieejām ir izveidot klasisku modeli, kas analizē šos parametrus un aprēķina sagaidāmo cenu (9.2.13. attēls). Šādai pieejai nepieciešama precīza un iepriekš zināma formula, kas reālajā praksē ir praktiski neiespējami.



**Attēls 9.2-13 Mājas vērtības novērtēšanai var izmantot klasisko algoritmu, kurā ir noteikta formula, pēc kuras jāatrod vērtība.**

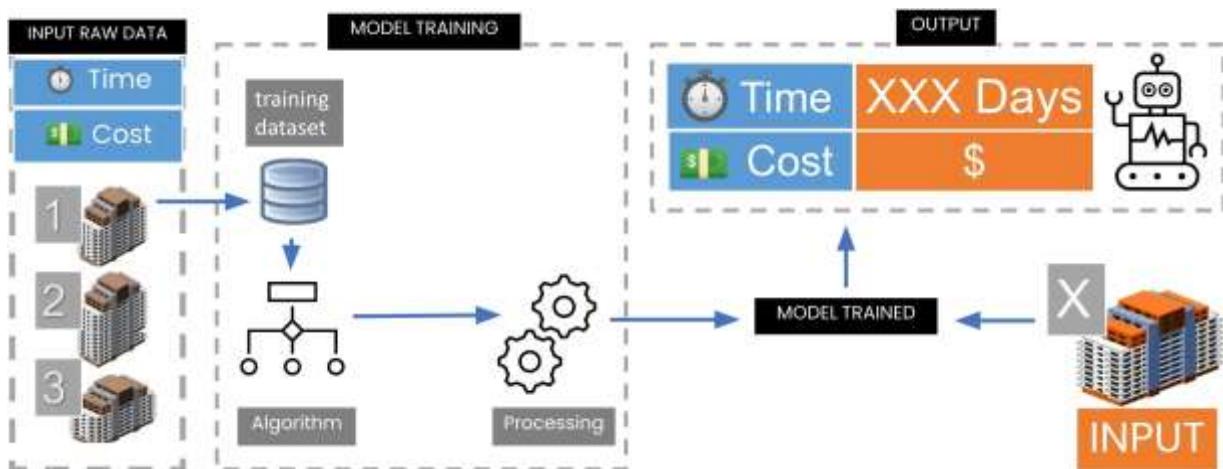
Mašīnmācīšanās novērš manuālu formulu meklēšanu un aizstāj tās ar apmācītiem algoritmiem, kas patstāvīgi nosaka atkarības, kuras ir daudzkārt precīzākas par iepriekš noteiktām vienādojumiem. Kā alternatīvu izveidosim mašīnmācīšanās algoritmu, kas radīs modeļi, pamatojoties uz iepriekšēju izpratni par problēmu un vēsturiskiem datiem, kas var būt nepilnīgi (9.2-14. attēls).

Izmantojot kā piemēru cenu noteikšanu, mašīnmācīšanās ļauj izveidot dažāda veida matemātiskos modeļus, kuriem nav nepieciešamas zināšanas par precīzu izmaksu veidošanās mehānismu. Modelis "mācās" no iepriekšējo projektu datiem, pielāgojoties reāliem likumsakarībām starp būvniecības parametriem, to izmaksām un termiņiem.



Attēls 9.2-14 Atšķirībā no klasiskās uz formulām balstītās aplēses mašīnmācīšanās algoritms tiek apmācīts, izmantojot vēsturiskos datus.

Uzraudzītas mašīnmācīšanās kontekstā, katrs projekts mācību datu kopā satur gan ievades atribūtus (piemēram, līdzīgu ēku izmaksas un laiku), gan paredzamās izejas vērtības (piemēram, izmaksas vai laiku). Līdzīgu datu kopu izmanto, lai izveidotu un pielāgotu mašīnmācīšanās modeli (9.2-15. attēls). Jo lielāka datu kopa un jo kvalitatīvāki tajā esošie dati, jo precīzāks būs modelis un precīzāki būs prognozēšanas rezultāti.



Attēls 9.2-15 ML modelis, kas apmācīts, izmantojot datus par izmaksām un grafiku no iepriekšējiem projektiem, ar noteiktu varbūtību noteiks jauna projekta izmaksas un grafiku.

Kad modelis ir izveidots un apmācīts, lai novērtētu jauna projekta būvniecību, vienkārši sniedziet modelim jaunus jaunā projekta atribūtus, un modelis ar zināmu varbūtību nodrošinās aplēstos rezultātus, pamatojoties uz iepriekš apgūtajiem modeļiem.

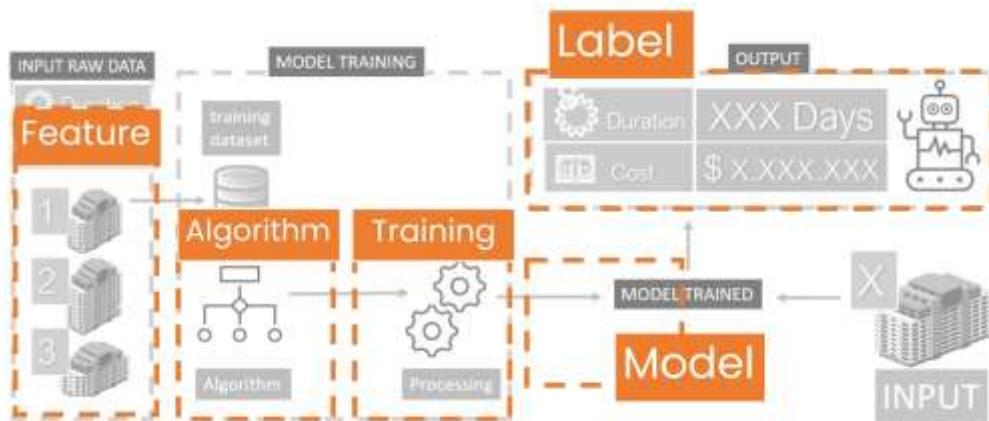
## Mašīnmācīšanās pamatlēdzieni

Mašīnmācīšanās nav maģija, tā ir tikai matemātika, dati un likumsakarību meklēšana. Tai nav īsta intelekta, bet tā ir programma, kas apmācīta uz datiem, lai atpazītu modeļus un pieņemtu lēmumus

bez pastāvīgas cilvēka iesaistīšanās.

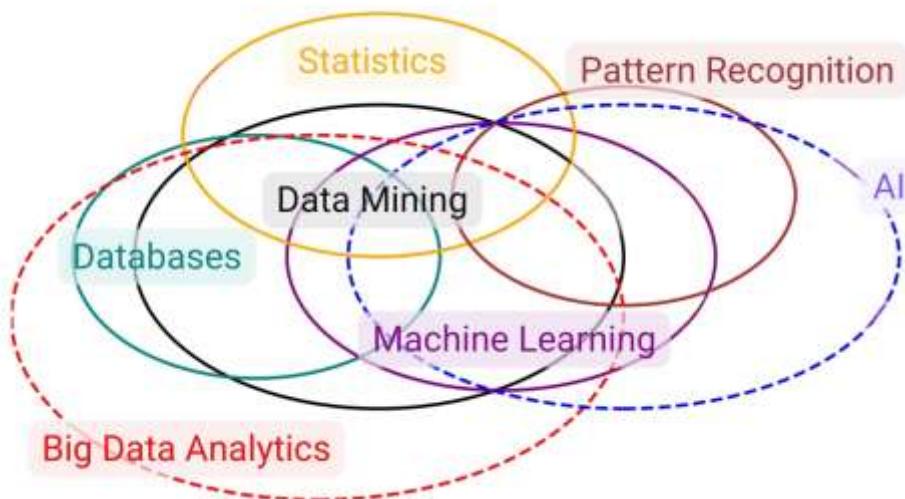
Mašīnmācīšanās struktūrai aprakstīt izmanto vairākus galvenos jēdzienus ( 9.2-16. attēls):

- **Etiķetes** ir mērķa mainīgie vai atribūti (parametrs "Izdzīvojušais" "Titānika" datu kopā), kas modelim jāparedz. Piemērs: būvniecības izmaksas (piemēram, dolāros), būvniecības darbu ilgums (piemēram, mēnešos).
- **Īpašības** ir neatkarīgie mainīgie vai atribūti, kas kalpo kā modeļa ievades dati. Prognozēšanas modelī tie tiek izmantoti, lai prognozētu marķējumus. Piemēri: zemes gabala lielums (kvadrātmetros), ēkas stāvu skaits, ēkas kopējā platība (kvadrātmetros), ģeogrāfiskā atrašanās vieta (ģeogrāfiskais platumis un garums), būvniecībā izmantoto materiālu veids. Raksturlielumu skaits nosaka arī datu dimensiju.
- **Modelis** ir dažādu hipotēžu kopums, no kurām viena aproksimē prognozējamo vai aproksimējamo mērķa funkciju. Piemērs: mašīnmācīšanās modelis, kas izmanto regresijas analīzes metodes, lai prognozētu būvniecības izmaksas un termiņus.
- **Mācīšanās algoritms** Mācīšanās algoritms ir process, kurā, izmantojot apmācības datu kopu, tiek atrasta vislabākā hipotēze modelī, kas precīzi atbilst mērķa funkcijai. Piemērs: lineārās regresijas, KNN vai izlases meža algoritms, kas analizē izmaksu un būvniecības laika datus, lai noteiktu sakarības un modeļus.
- **Apmācīšana** - apmācības procesa laikā algoritms analizē apmācības datus, lai atrastu modeļus, kas atbilst attiecībām starp ievades atribūtiem un mērķa marķējumiem. Šā procesa rezultāts ir apmācis mašīnmācīšanās modelis, kas ir gatavs prognozēšanai. Piemērs: process, kurā algoritms analizē vēsturiskus būvniecības datus (izmaksas, laiks, objekta raksturlielumi), lai izveidotu prognozēšanas modeli.



Attēls 9.2-16 ML izmanto marķējumus un atribūtus, lai izveidotu modeļus, kas tiek apmācīti, izmantojot algoritmus, lai prognozētu rezultātus.

Mašīnmācīšanās nepastāv izolēti, bet ir daļa no plašākas analītisko disciplīnu ekosistēmas, tostarp statistikas, datu bāzu, datu ieguves, tēlu atpazīšanas, lielo datu analīzes un mākslīgā intelekta. Attēlā 9.2-17 parādīts, kā šīs jomas pārklājas un papildina viena otru, nodrošinot visaptverošu sistēmu modernām lēmumu pieņemšanas un automatizācijas sistēmām.



Attēls 9.2-17 Attiecības starp dažādām datu analīzes jomām: statistiku, mašīnmācīšanos, mākslīgo intelektu, lielajiem datiem, attēlu atpazīšanu un datu ieguvi.

Mašīnmācīšanās galvenais mērķis ir piešķirt datoriem spēju automātiski apgūt zināšanas bez cilvēka iejaukšanās vai palīdzības un attiecīgi pielāgot savu darbību [159].

Tātad nākotnē cilvēka loma būs tikai nodrošināt mašīnai izziņas spējas - viņš noteiks nosacījumus, svarus un parametrus, un mašīnmācīšanās modelis paveiks visu pārējo.

Nākamajā nodaļā mēs aplūkosim konkrētus algoritmu pielietošanas piemērus. Tiks izmantotas reālas tabulas un vienkāršoti modeļi, lai parādītu, kā soli pa solim tiek veidota prognoze.



## NODAĻA 9.3.

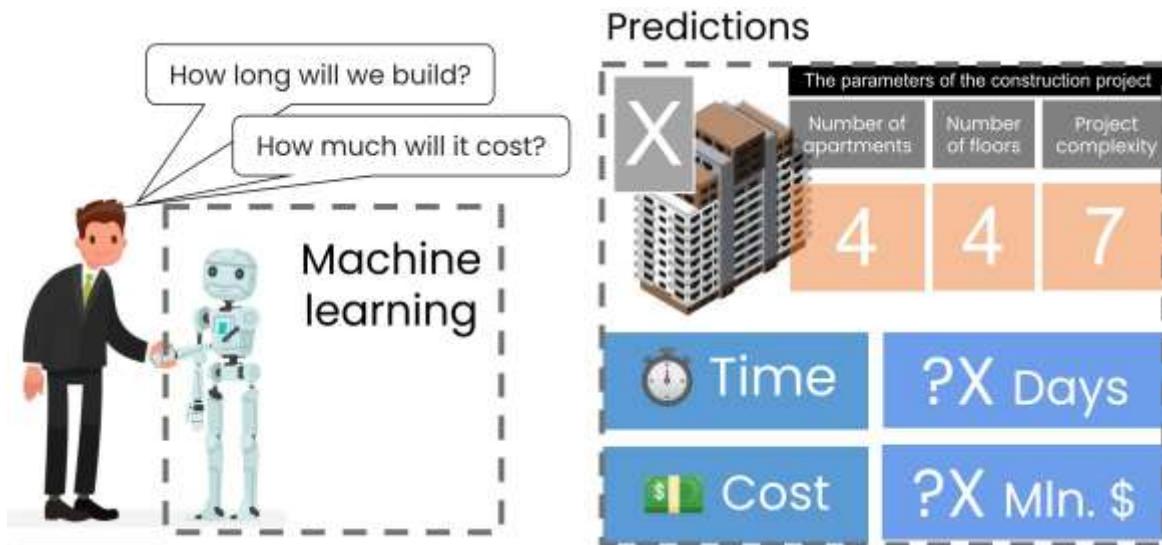
### IZMAKSU UN TERMIŅU PROGNOZĒŠANA, IZMANTOJOT MAŠĪNMĀCĪŠANOS

**Piemērs mašīnmācīšanās izmantošanai, lai atrastu projekta izmaksas un termiņus.**

Būvdarbu laika un izmaksu aplēse ir viens no galvenajiem procesiem būvniecības uzņēmuma darbībā. Tradicionāli šādus aprēķinus veic eksperti, pamatojoties uz pieredzi, uzziņu grāmatām un normatīvo aktu datubāzēm. Tomēr līdz ar digitālo transformāciju un datu pieejamības palielināšanos tagad ir iespējams izmantot mašīnmācīšanās (ML) modeļus, lai uzlabotu šādu aplēšu precizitāti un automatizētu to veikšanu.

Mašīnmācīšanās ieviešana būvniecības izmaksu un termiņu aprēķināšanas procesā ne tikai ļauj efektīvāk plānot, bet arī kļūst par sākumpunktu inteliģento modeļu integrēšanai citos uzņēmējdarbības procesos - sākot no riska pārvaldības līdz loģistikas un iepirkumu optimizācijai.

Ir svarīgi ātri noteikt, cik ilgā laikā būs jābūvē projekts un kādas būs tā kopējās izmaksas. Šie jautājumi par projekta laiku un izmaksām tradicionāli ir bijuši gan klientu, gan būvniecības uzņēmumu uzmanības centrā jau kopš būvniecības nozares pirmsākumiem.



Attēls 9.3-1 Būvniecības projektos būvniecības laika un izmaksu aplēses ātrums un kvalitāte ir galvenie veiksmes faktori.

Turpmākajā piemērā tiks iegūti galvenie dati no iepriekšējiem projektiem un izmantoti mašīnmācīšanās modeļa izstrādei, kas ļaus mums izmantot modeli, lai novērtētu jaunu būvniecības projektu izmaksas un termiņus ar jauniem parametriem (9.3-1. attēls).

Aplūkojiet trīs projektus ar trim galvenajiem raksturlielumiem: dzīvokļu skaits (kur 100 dzīvokļi ir

līdzvērtīgi skaitlim 10, lai atvieglotu vizualizāciju), stāvu skaits un nosacīts būvniecības sarežģītības rādītājs skalā no 1 līdz 10, kur 10 ir augstākais sarežģītības rādītājs. Mašīnmācīšanās procesā vērtību, piemēram, 100 pārvēršanu 10 vai 50 - 5, sauc par "normalizāciju".

**Normalizācija** mašīnmācībā ir process, kurā dažādi skaitliskie dati tiek saskaņoti kopīgā mērogā, lai atvieglotu apstrādi un analīzi. Šis process ir īpaši svarīgs, ja datiem ir dažādi mērogi un mērvienības.

Pieņemsim, ka pirmajā projektā ( 9.3-2. attēls) bija 50 dzīvokļi (pēc normalizācijas - 5), 7 stāvi un sarežģītības rādītājs 2, kas nozīmē, ka būvniecība ir salīdzinoši vienkārša. Otrajā projektā jau bija 80 dzīvokļi, 9 stāvi un salīdzinoši sarežģīta konstrukcija. Šādos apstākļos pirmās un otrās daudzdzīvokļu ēkas būvniecība ilga 270 un 330 dienas, un projekta kopējās izmaksas bija attiecīgi 4,5 miljoni un 5,8 miljoni ASV dolāru.

Construction project			The key parameters of the project		
	Number of apartment	Number of floors	Project complexity	Time	Cost
1	5	7	2	270	\$ 4.502.000
2	8	9	6	330	\$ 5.750.000
3	3	5	3	230	\$ 3.262.000
X	4	4	7	?X	\$ ?X. XXX.XXX

Attēls 9.3-2 leprieķējo projektu kopuma piemērs, kas tiks izmantots, lai novērtētu nākamā projekta X laiku un izmaksas.

Veidojot mašīnmācīšanās modeli šādiem datiem, galvenais uzdevums ir noteikt prognozēšanai kritiskos atribūtus (vai etiķetes), šajā gadījumā - būvniecības laiku un izmaksas. Izmantojot nelielu datu kopu, mēs izmantojam informāciju par iepriekšējiem būvniecības projektiem, lai plānotu jaunus projektus: izmantojot mašīnmācīšanās algoritmus, mums jāparedz jauna projekta X izmaksas un būvniecības ilgums, pamatojoties uz dotajiem jaunā projekta atrībūtiem, piemēram, 40 dzīvokļi, 4 stāvi un relatīvi augsta projekta sarežģītība 7 ( 9.3-2. attēls). Reālā vidē ievades parametru skaits var būt daudz lielāks, sākot no vairākiem desmitiem līdz pat simtiem faktoru. Tie var būt šādi: būvmateriālu veids, klimatiskā zona, būvuzņēmēju kvalifikācijas līmenis, inženierkomunikāciju pieejamība, pamatu veids, darbu uzsākšanas sezona, meistaru komentāri utt.

Lai izveidotu prognozēšanas mašīnmācīšanās modeli, mums ir jāizvēlas algoritms tā izveidei. Algoritms mašīnmācībā ir kā matemātiska recepte, kas māca datoru, kā veikt prognozes (sajaukt pareizajā secībā parametrus) vai pieņemt lēmumus, pamatojoties uz datiem.

Lai analizētu datus par iepriekšējiem būvniecības projektiem un prognozētu turpmāko projektu laiku un izmaksas (9.3-2. attēls), var izmantot vienu populāru mašīnmācīšanās algoritmu:

- **Lineārā regresija (Lineārā regresija):** ar šo algoritmu mēģina atrast tiešu saistību starp atribūtiem, piemēram, starp stāvu skaitu un būvniecības izmaksām. Algoritma mērķis ir atrast lineāru vienādojumu, kas vislabāk apraksta šo sakarību un ļauj veikt prognozes.
- **Algoritms k-nearest neighbours (k-NNN):** šis algoritms salīdzina jaunu projektu ar iepriekšējiem projektiem, kas bija līdzīgi pēc lieluma vai sarežģītības. K-NN klasificē datus, pamatojoties uz to, kuri no k (skaits) mācību piemēriem ir vistuvāki. Regresijas kontekstā rezultāts ir k tuvāko kaimiņu vidējā vērtība vai mediāna.
- **Lēmumu koki:** ir prognozēšanas modelēšanas modelis, kas, izmantojot koku struktūru, sadala datus apakškopās, pamatojoties uz dažādiem nosacījumiem. Katrs koka mezglis ir nosacījums vai jautājums, kas novēr pie tālākas datu dalīšanas, un katras lapa ir galīgais paredzējums vai iznākums. Algoritms sadala datus mazākās grupās, pamatojoties uz dažādām īpašībām, piemēram, vispirms pēc stāstu skaita, pēc tam pēc sarežģītības utt. un tā tālāk, lai veiktu prognozēšanu.

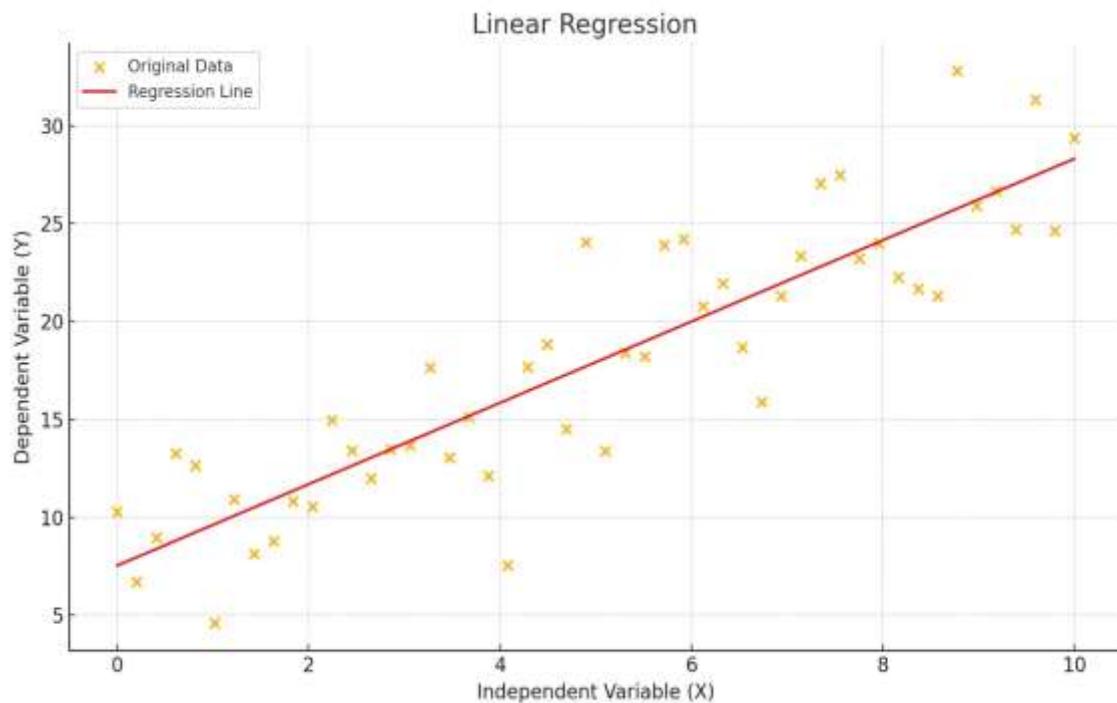
Aplūkosim mašīnmācīšanās algoritmus jauna projekta izmaksu aplēsēm, kā piemērus izmantojot divus populārus algoritmus: lineāro regresiju un K-nearest neighbours algoritmu.

## Projekta izmaksu un laika prognozēšana, izmantojot lineāro regresiju

Lineārā regresija ir fundamentāls datu analīzes algoritms, kas prognozē mainīgā lieluma vērtību, pamatojoties uz lineāru saistību ar vienu vai vairākiem citiem mainīgajiem lielumiem. Šajā modelī tiek pieņemts, ka pastāv tieša lineāra sakarība starp atkarīgo mainīgo un vienu vai vairākiem neatkarīgajiem mainīgajiem, un algoritma mērķis ir atrast šo sakarību.

Lineārās regresijas vienkāršība un skaidrība ir padarījusi to par populāru instrumentu dažādās jomās. Ja runa ir par vienu mainīgo lielumu, lineārās regresijas mērķis ir atrast vispiemērotāko līniju, kas iet caur datu punktiem.

Lineārā regresija atrod vislabāko taisni (sarkanā līnija), kas aproksimē atkarību starp ieejas mainīgo X un izejas mainīgo Y. Šī līnija ļauj prognozēt Y vērtības jaunām X vērtībām, pamatojoties uz identificēto lineāro sakarību (9.3-3. attēls).



Attēls 9.3-3 Lineārās regresijas princips ir atrast vislabāko taisni caur mācību vērtībām.

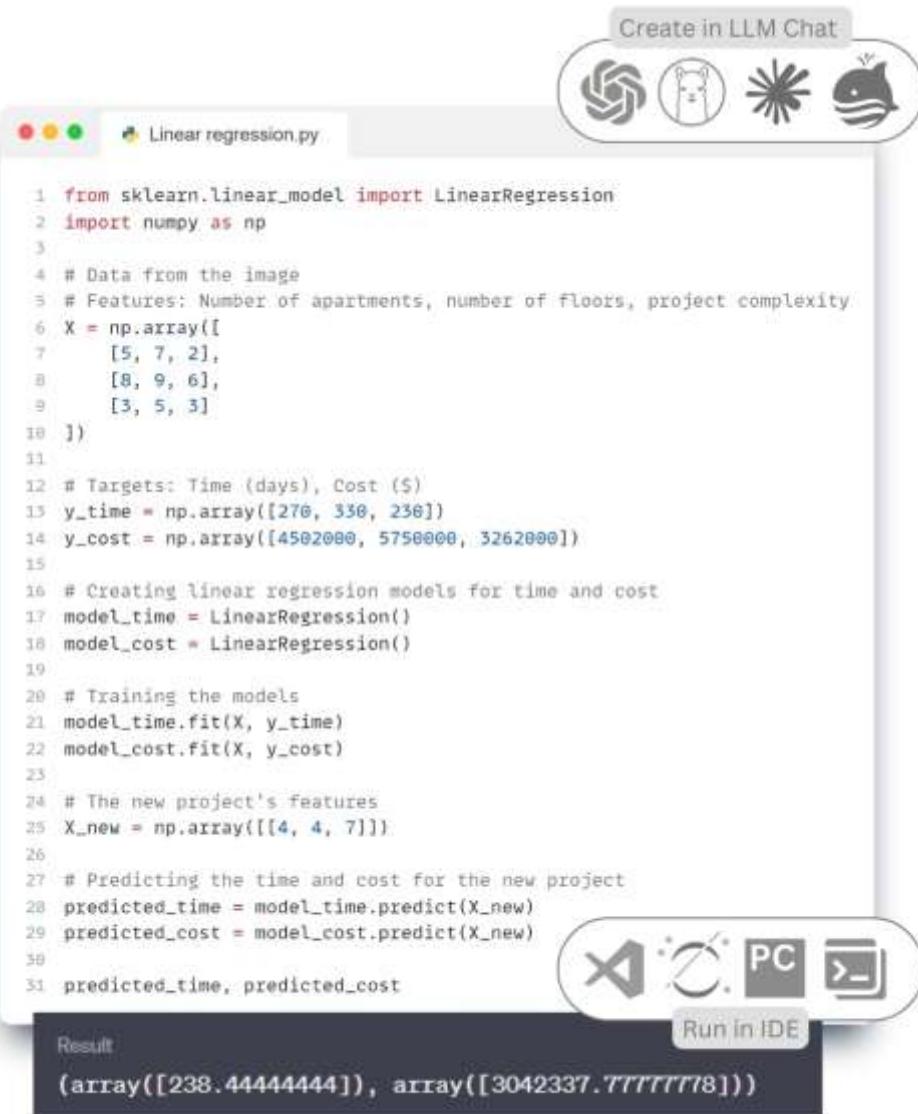
Šo rindu attēlo vienādojums, kurā, ievadot neatkarīgā mainīgā lieluma ( $X$ ) vērtību, iegūst prognozēto atkarīgā mainīgā lieluma ( $Y$ ) vērtību. Šis process ļauj efektīvi prognozēt  $Y$  no zināmām  $X$  vērtībām, izmantojot lineāro sakarību starp tām. Piemērs šādas statistiski vidējās taisnes atrašanai redzams Sanfrancisko būvatļauju datu novērtējumā (9.1-7. attēls), kur inflācija, izmantojot lineāro regresiju, tika aprēķināta dažādiem objektu veidiem.

Ielādēsim projekta datu tabulu (9.3-2. attēls no iepriekšējās nodaļas) tieši LLM un lūgsim to izveidot mums vienkāršu mašīnmācīšanās modeli.

- ⌚ Nosūtīt teksta pieprasījumu LLM tērzēšanai (CHATGP, LlaMa, Mistral DeepSeek, Grok, Claude, QWEN):

Mums ir jāparāda vienkārša mašīnmācīšanās modeļa uzbūve, lai prognozētu jauna projekta  $X$  realizācijas izmaksas un laiku (9.3-2. attēls kā pievienotais attēls). ↴

 LLM atbilde :



```

1  from sklearn.linear_model import LinearRegression
2  import numpy as np
3
4  # Data from the image
5  # Features: Number of apartments, number of floors, project complexity
6  X = np.array([
7      [5, 7, 2],
8      [8, 9, 6],
9      [3, 5, 3]
10 ])
11
12 # Targets: Time (days), Cost ($)
13 y_time = np.array([270, 330, 230])
14 y_cost = np.array([4502000, 5750000, 3262000])
15
16 # Creating linear regression models for time and cost
17 model_time = LinearRegression()
18 model_cost = LinearRegression()
19
20 # Training the models
21 model_time.fit(X, y_time)
22 model_cost.fit(X, y_cost)
23
24 # The new project's features
25 X_new = np.array([[4, 4, 7]])
26
27 # Predicting the time and cost for the new project
28 predicted_time = model_time.predict(X_new)
29 predicted_cost = model_cost.predict(X_new)
30
31 predicted_time, predicted_cost

```

Result  
`(array([238.44444444]), array([3042337.77777778]))`

Attēls 9.3-4 LLM izvēlējās lineāro regresiju, lai izveidotu mašīnmācīšanās modeli projekta izmaksu un laika prognozēšanai.

LLM automātiski atpazina tabulu no pievienotā attēla un konvertēja datus no vizuālā formāta uz tabulas masīvu ( 9.3-4. attēls - 6. rinda). Šo masīvu izmantoja par pamatu pazīmju un marķējumu izveidei, no kuriem tika izveidots mašīnmācīšanās modelis ( 9.3-4. attēls - 17.-22. rinda), kurā izmantoja lineāro regresiju.

Izmantojot pamata lineārās regresijas modeli, kas tika apmācis, izmantojot "ārkārtīgi mazu" datu kopu, tika veiktas prognozes jaunam hipotētiskam būvniecības projektam ar nosaukumu X projekts.

Mūsu uzdevumā šim projektam ir 40 dzīvokļi, 4 stāvi un 7 sarežģītības līmenis ( 9.3-2. attēls).

Kā prognozēts lineārās regresijas modelī, kas balstīts uz ierobežotu un nelielu datu kopu jaunajam X projektam ( 9.3-4. attēls - 24.-29. līnija):

- **Būvniecības ilgums** būs aptuveni 238 dienas (238.4444444).
- **Kopējie izdevumi** būs aptuveni \$ 3 042 338 (3042337,777).

Lai turpinātu pētīt projekta izmaksu hipotēzi, ir lietderīgi eksperimentēt ar dažādiem mašīnmācīšanās algoritmiem un metodēm. Tāpēc prognozēsim tādas pašas izmaksu un laika vērtības jaunam projektam X, pamatojoties uz nelielu vēsturisko datu kopu, izmantojot K-tuvāko kaimiņu algoritmu (k-NN).

### **Projekta izmaksu un laika prognozes, izmantojot K-tuvāko kaimiņu algoritmu (k-NN).**

Mēs izmantojam k-Nearest Neighbours (k-NN) algoritmu kā papildu prognozētāju, lai novērtētu jauna projekta izmaksas un ilgumu. K-tuvāko kaimiņu (k-NN) algoritms ir uzraudzīta mašīnmācīšanās (supervised machine learning) metode gan klasifikācijai, gan regresijai. Mēs jau iepriekš esam apsprieduši k-NN algoritmu arī vektoru datubāzes meklēšanas kontekstā ( 8.2-2. attēls), kur to izmanto, lai atrastu tuvākos vektorus (piemēram, tekstu, attēlus vai tehniskos aprakstus). Šajā pieejā katrs projekts tiek attēlots kā punkts daudzdimensiju telpā, kur katra dimensija atbilst konkrētam projekta atribūtam.

Mūsu gadījumā, nemot vērā katras projekta trīs atribūtus, mēs tos attēlosim kā punktus trīsdimensiju telpā ( 9.3-5. attēls). Tādējādi mūsu gaidāmais projekts X tiks lokalizēts šajā telpā ar koordinātām ( $x=4$ ,  $y=4$ ,  $z=7$ ). Jāatzīmē, ka reālos apstākļos punktu skaits un telpas dimensionalitāte var būt par kārtām lielāka.

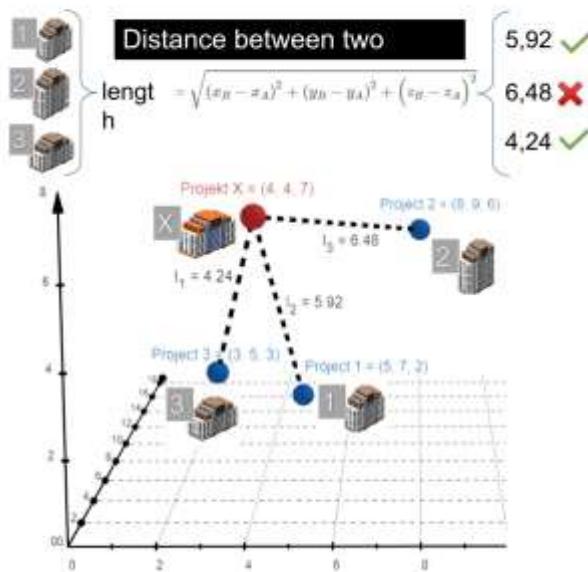
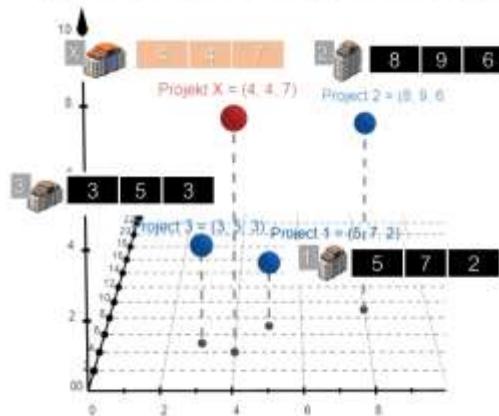
K-NN (k-nearest neighbours) algoritms darbojas, mērot attālumu starp vēlamo projektu X un projektiem mācību datu bāzē. Salīdzinot šos attālumus, algoritms nosaka projektus, kas ir vistuvāk jaunā projekta X punktam.

Piemēram, ja otrs projekts ( $x=8$ ,  $y=9$ ,  $z=6$ ) no mūsu sākotnējās datu kopas ir daudz tālāk no X ( 9.3-5. attēls) nekā citi projekti, to var izslēgt no turpmākās analīzes. Rezultātā aprēķinos var izmantot tikai divus ( $k=2$ ) tuvākos projektus, uz kuru pamata tiks noteikta vidējā vērtība.

Šāda metode, izmantojot kaimiņu meklēšanu, ļauj novērtēt līdzības starp projektiem, kas savukārt palīdz izdarīt secinājumus par jauna projekta iespējamām izmaksām un termiņiem, pamatojoties uz līdzīgiem iepriekš īstenotiem projektiem.

## k-nearest neighbors algorithm

The KNN algorithm assumes that similar things exist in close proximity. In other words, similar things are near to each other.



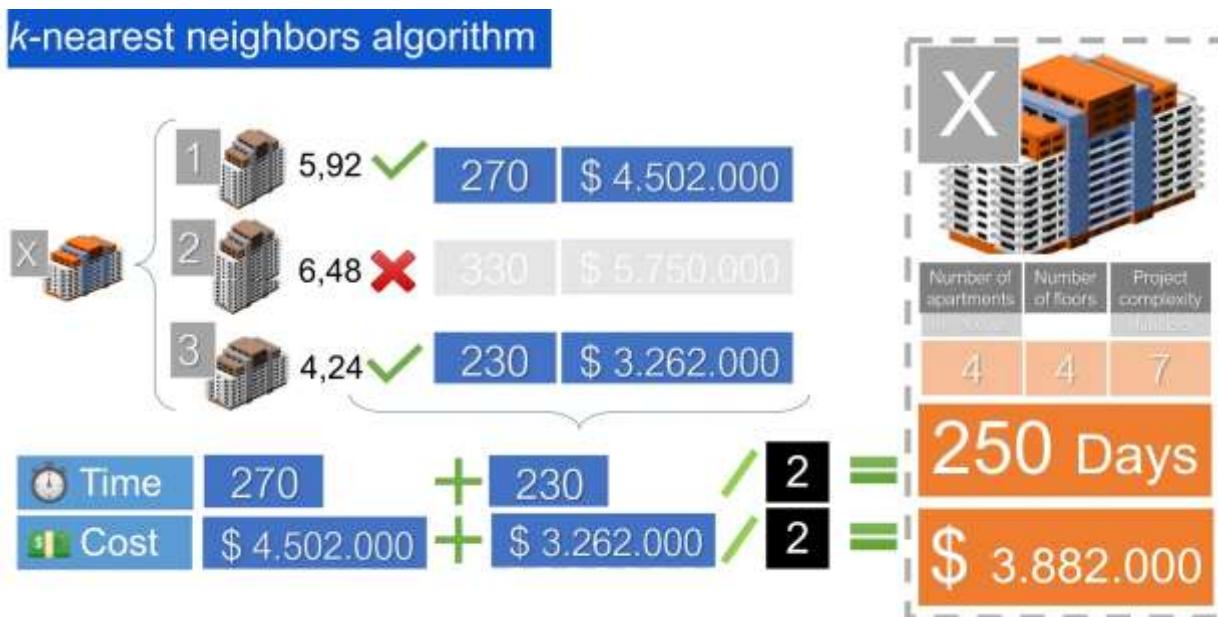
Attēlā 9.3-5 K-NN algoritmā projekti tiek attēloti kā punkti daudzdimensiju telpā, un tuvākie projekti tiek atlasīti, pamatojoties uz attālumiem, lai novērtētu un prognozētu līdzību.

K-NN darbs ietver vairākus galvenos soļus:

- **Datu sagatavošana:** vispirms tiek ielādētas mācību un testa datu kopas. Mācību datus izmanto algoritma "apmācībai", bet testa datus izmanto, lai pārbaudītu tā efektivitāti.
- **Parametra K izvēle:** tiek izvēlēts skaitlis K, kas norāda, cik tuvākie kaimiņi (datu punkti) jāņem vērā algoritmā. "K" vērtība ir ļoti svarīga, jo tā ietekmē rezultātu.
- Klasifikācijas process un regresija testa datiem:
  - Attālumu **aprēķināšana:** katram elementam no testa datiem tiek aprēķināts attālums līdz katram elementam no mācību datiem ( 9.3-5. attēls). Šim nolūkam var izmantot dažādas attāluma mērišanas metodes, piemēram, Eiklīda attālumu (visizplatītākā metode), Manhetenas attālumu vai Haminga attālumu.
  - **K tuvāko kaimiņu šķirošana un atlase:** pēc attālumu aprēķināšanas tie tiek sašķiroti un atlasīti K tuvākie punkti testa punktam.
  - **Testa punkta klases vai vērtības noteikšana:** ja tas ir klasifikācijas uzdevums, testa punkta klasi nosaka, pamatojoties uz visbiežāk sastopamo klasi starp K izvēlētajiem kaimiņiem. Ja tas ir regresijas uzdevums, aprēķina K kaimiņu vērtību vidējo vērtību (vai citu centrālās tendences rādītāju).
- **Procesa pabeigšana:** kad visi testa dati ir klasificēti vai tiem ir veiktas prognozes, process ir pabeigts.

Algoritms k-nearest neighbours (k-NN) ir efektīvs daudzos praktiskos lietojumos un ir viens no galvenajiem mašīnmācīšanās speciālistu arsenāla rīkiem. Šis algoritms ir populārs, pateicoties tā vienkāršbai un efektivitātei, īpaši uzdevumos, kur attiecības starp datiem ir viegli interpretējamas.

Mūsu piemērā pēc K tuvāko kaimiņu algoritma piemērošanas tika identificēti divi projekti (no mūsu nelielās izlases) ar mazāko attālumu līdz projektam X (attēls 9.3-5). Pamatojoties uz šiem projektiem, algoritms nosaka to vidējo cenu un būvniecības ilgumu. Pēc analīzes (9.3-6. attēls) algoritms, vidēji aprēķinot tuvāko kaimiņu rādītājus, secina, ka projekts X maksās aptuveni \$ 3 800 000 USD un tā pabeigšana prasīs aptuveni 250 dienas.



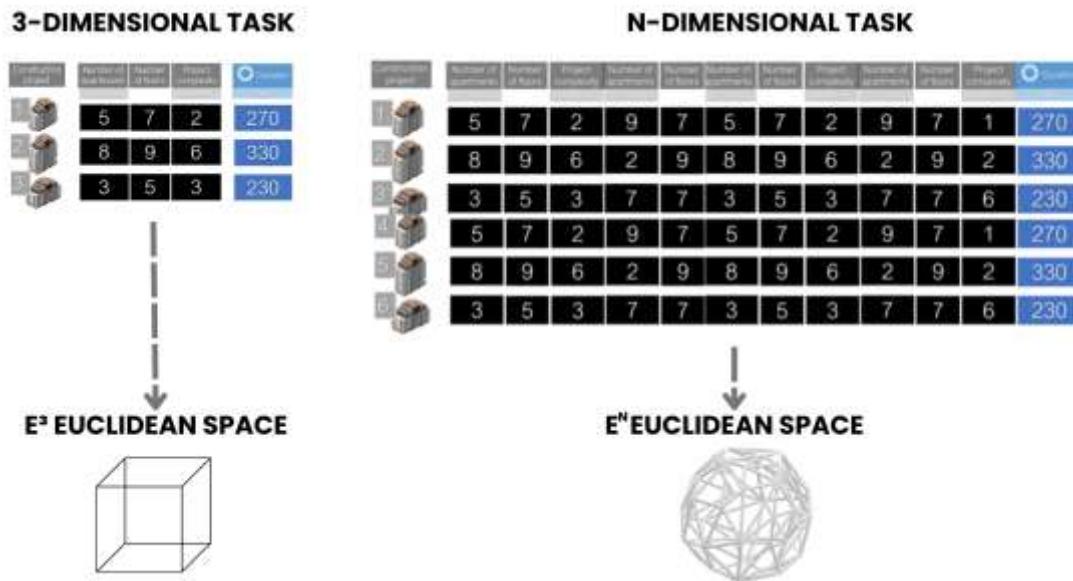
Attēls 9.3-6 K-nearest neighbours algoritms nosaka projekta X izmaksas un grafiku, analizējot divus tuvākos projektus izlasē.

Algoritms k-Nearest Neighbors (k-NN) ir īpaši populārs klasifikācijas un regresijas uzdevumos, piemēram, ieteikumu sistēmās, kur to izmanto, lai ieteiktu produktus vai saturu, pamatojoties uz preferencēm, kas ir līdzīgas konkrēta lietotāja interesēm. Turklāt k-NN tiek plaši izmantots medicīnās diagnostikā, lai klasificētu slimību veidus, pamatojoties uz pacientu simptomiem, modeļu atpazīšanā un finanšu nozarē, lai novērtētu klientu kredītspēju.

Pat ar ierobežotiem datiem mašīnmācīšanās modeļi var sniegt noderīgas prognozes un būtiski uzlabot būvniecības projektu vadības analītisko komponentu. Paplašinot un attīrot vēsturiskos datus, ir iespējams pāriet uz sarežģītākiem modeļiem - piemēram, nemot vērā būvniecības veidu, vietu, būvniecības uzsākšanas sezonu un citus faktorus.

Mūsu vienkāršotajā uzdevumā vizualizācijai 3D telpā tika izmantoti trīs atribūti, taču reālajos projektos

vidēji ir simtiem vai tūkstošiem atribūtu (sk. datu kopu no nodaļas "Uz CAD balstītu lielo datu piemērs (BIM)", kas ievērojami palielina telpas dimensionalitāti un projektu attēlošanas sarežģītību vektoru veidā (9.3-7. attēls).



Attēls 9.3-7 Vienkāršotajā piemērā 3D -vizualizācijai tika izmantoti trīs atribūti, bet reālos projektos to ir vairāk.

Piemērojot dažādus algoritmus vienai un tai pašai datu kopai projektam X, kurā ir 40 dzīvokļi, 4 stāvi un sarežģītības līmenis 7, tika iegūtas dažādas prognozētās vērtības. Ar lineārās regresijas algoritmu tika prognozēts pabeigšanas laiks 238 dienas un izmaksas \$ 3 042 338 (9.3-4. attēls), savukārt ar k-NN algoritmu tika prognozētas 250 dienas un \$ 3 882 000 (9.3-6. attēls).

Ar mašīnmācīšanās modeļiem iegūto prognožu precizitāte, ir tieši atkarīga no ievades datu apjoma un kvalitātes. Jo vairāk projektu ir iesaistīti apmācībā un jo pilnīgāk un precīzāk ir attēlotas to īpašības (atribūti) un rezultāti (etiķetes), jo lielāka varbūtība iegūt ticamus prognozes ar minimālām kļūdu vērtībām.

Šajā procesā liela nozīme ir datu pirmapstrādes metodēm, tostarp:

- Normalizācija, lai funkcijas saskaņotu ar vienotu skalu;
- Noviržu noteikšana un novēršana, novēršot modeļa izkroplojumus;
- Kategorīalu atribūtu kodēšana, lai varētu manipulēt ar teksta datiem;
- Trūkstošo vērtību aizpildīšana, palielinot modeļa robustumu.

Turklāt, lai novērtētu modeļa vispārināmību un tā noturību pret jaunām datu kopām, tiek izmantotas savstarpējas apstiprināšanas metodes, lai atklātu pārmēriju pielāgošanu un uzlabotu prognozēšanas ticamību.

*Haoss ir atšifrējama kārtība [160].*

- Hosē Saramago, "Dubultnieks"

Pat ja jums šķiet, ka jūsu uzdevumu haosu nevar aprakstīt formāli, jums jāzina, ka jebkurš notikums pasaulē un jo īpaši būvniecības procesi ir pakļauti matemātiskiem likumiem, kuriem var būt nepieciešams atbalsts, aprēķinot vērtības nevis ar stingru formulu palīdzību, bet gan ar statistikas un vēsturisko datu palīdzību.

Gan tradicionālās izmaksu aprēķinu nodaļas aplēses, gan mašīnmācīšanās modeļi neizbēgami saskaras ar nenoteiktību un iespējamiem kļūdu avotiem. Tomēr, ja ir pieejami pietiekami kvalitatīvi dati, mašīnmācīšanās modeļi var uzrādīt salīdzināmu un dažkārt pat augstāku prognozēšanas precizitāti nekā ekspertu vērtējumi.

Mašīnmācīšanās, visticamāk, kļūs par uzticamu papildu rīku analīzēm, kas var: precizēt aprēķinus, ierosināt alternatīvus scenārijus un noteikt slēptās atkarības starp projekta parametriem. Šādi modeļi nepretendēs uz universālām iespējām, taču drīz tie ieņems nozīmīgu vietu aprēķinos un lēmumu pieņemšanas procesos. Mašīnmācīšanās tehnoloģijas neizslēgs inženieru, tāmētāju un analītiķu līdzdalību, bet gluži pretēji - paplašinās viņu iespējas, piedāvājot papildu skatpunktu, kas balstīts uz vēsturiskiem datiem.

Ja mašīnmācīšanās tiks pareizi integrēta būvniecības uzņēmumu biznesa procesos, tai ir potenciāls kļūt par svarīgu vadības lēmumu atbalsta sistēmas elementu - nevis kā cilvēku aizstājējam, bet gan kā viņu profesionālās intuīcijas un inženiertehniskās loģikas papildinājumam.

## Nākamie soļi: no uzglabāšanas līdz analīzei un prognozēšanai

Mūsdienīgas pieejas darbam ar datiem sāk mainīt veidu, kā būvniecības nozarē tiek pieņemti lēmumi. Pāreja no intuitīviem novērtējumiem uz objektīvu datu analīzi ne tikai uzlabo precizitāti, bet arī paver jaunas iespējas optimizēt procesus. Apkopojoši šo daļu, ir vērts izcelt galvenos praktiskos soļus, kas palīdzēs jums piemērot aplūkotās metodes ikdienas uzdevumos:

- Ilgtspējīgas uzglabāšanas infrastruktūras izveide
  - Mēģiniet apvienot dažādus dokumentus un projekta datus vienotā tabulas modelī, apkopojojot galveno informāciju vienā datu rāmī turpmākai analīzei.
  - izmantot efektīvus datu glabāšanas formātus, piemēram, CSV vai XLSX vietā izmantot kolonnveida formātus, piemēram, Apache Parquet, jo īpaši tām kopām, kuras nākotnē varētu izmantot mašīnmācīšanās modeļu apmācībai.
  - Izveidot datu versiju sistēmu, lai sekotu līdzi izmaiņām visā projekta gaitā.
- Analīzes un automatizācijas rīku ieviešana

- Sākt analizēt vēsturiskos projektu datus - dokumentāciju, modeļus, aplēses - lai noteiktu modeļus, tendences un anomālijas.
- Apgūt ETL procesus (ekstrakts, pārveidošana, ielāde), lai automātiski ielādētu un sagatavotu datus.
- Uzziniet, kā vizualizēt galvenos rādītājus, izmantojot dažādas bezmaksas Python vizualizācijas bibliotēkas.
- sākt izmantot statistikas metodes un izlases veida paraugu ņemšanu, lai iegūtu reprezentatīvus un reproducējamus analītiskos rezultātus.

#### ■ Pieaug briedums darbā ar datiem

- Apgūstiet dažus pamata mašīnmācīšanās algoritmus, izmantojot vienkāršus un vienkāršus piemērus, piemēram, "Titānika" datu kopu.
- Analizējiet pašreizējos procesus un noskaidrojet, kurās jomās varat pāriet no stingras cēloņu un seku loģikas uz statistiskām prognozēšanas un novērtēšanas metodēm.
- Sākt uzskatīt datus par stratēģisku aktīvu, nevis blakusproduktu: veidot lēmumu pieņemšanas procesus, pamatojoties uz datu modeļiem, nevis uz konkrētiem programmatūras risinājumiem.

Būvniecības uzņēmumi, kas ir apzinājušies datu vērtību, sāk jaunu attīstības posmu, kurā konkurences priekšrocības nosaka nevis resursu apjoms, bet gan uz analītikas datiem balstītu lēmumu pieņemšanas ātrums.



## MAKSIMĀLI ĒRTA DRUKĀTĀ VERSIJA

Jūsu rīcībā ir bezmaksas digitālā versija **Data-Driven Construction**. Ērtākam darbam un ātrākai piekļuvei materiāliem iesakām pievērst uzmanību **drukātajam izdevumam**:



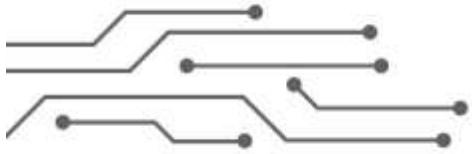
**Vienmēr pa rokai:** grāmata drukātā formātā kļūs par uzticamu darba rīku, kas ļaus ātri atrast un izmantot vajadzīgās vizualizācijas un diagrammas jebkurā darba situācijā.

**Augstas kvalitātes ilustrācijas:** visi attēli un grafikas drukātajā izdevumā ir maksimāli kvalitatīvi.

**Ātra piekļuve informācijai:** vienkārša navigācija, iespēja veikt piezīmes, grāmatzīmes un strādāt ar grāmatu jebkurā vietā.

legādājoties pilnu drukāto grāmatas versiju, jūs iegūstat ērtu rīku ērtam un efektīvam darbam ar informāciju: iespēju ātri izmantot vizuālos materiālus ikdienas uzdevumos, ātri atrast nepieciešamās shēmas un veikt piezīmes. Turklāt ar savu pirkumu atbalstāt atvērto zināšanu izplatīšanu.

Grāmatas drukāto versiju var pasūtīt: [datadrivenconstruction.io/books](http://datadrivenconstruction.io/books)



## X DAĻA

# **BŪVΝIECĪBAS NOZARE DIGITĀLAJĀ LAIKMETĀ. IESPĒJAS UN IZAICINĀJUMI**

Noslēdzošā, desmitā daļa ir visaptverošs ieskats būvniecības nozares nākotnē digitālās transformācijas laikmetā. Tajā analizēta pāreja no cēloņsakarību analīzes uz darbu ar lielo datu kopsakarībām. Tieki vilktas paralēles starp tēlotājmākslas attīstību un datu darba attīstību būvniecībā, parādot, kā nozare virzās no detalizētas kontroles uz holistisku izpratni par procesiem. Tieki pētīta būvniecības nozares "uberizācijas koncepcija", kur datu pārredzamība un automatizēti aprēķini var radikāli mainīt tradicionālos uzņēmējdarbības modeļus, novēršot nepieciešamību pēc starpniekiem un samazinot spekulāciju iespējas. Detalizēti tiek aplūkoti neatrisinātie jautājumi, piemēram, elementu universālā klasifikācija, dodot būvniecības uzņēmumiem laiku pielāgoties jaunajai videi. Daļa noslēdzas ar konkrētiem ieteikumiem digitālās transformācijas stratēģijas veidošanai, kas ietver ievainojamību analīzi un pakalpojumu paplašināšanu, lai saglabātu konkurētspēju mainīgajā nozarē.

## NODĀLA 10.1.

### IZDZĪVOŠANAS STRATĒGIJAS: KONKURENCES PRIEKŠROCĪBU VEIDOŠANA

#### Korelācijas aprēķinu vietā: būvniecības analītikas nākotne

Pateicoties straujai informācijas digitalizācijai (1.1.1-5. attēls), mūsdienu būvniecībā notiek fundamentālas pārmaiņas, un dati kļūst ne tikai par rīku, bet stratēģisku aktīvu, kas var būtiski mainīt tradicionālo pieeju projektu un uzņēmējdarbības vadībai.

Tūkstošiem gadu būvniecībā ir izmantotas deterministiskas metodes - precīzi aprēķini, detalizācija un stingra parametru kontrole. Pirmajos mūsu ēras gadsimtos romiešu inženieri izmantoja matemātiskos principus akveduktu un tiltu būvniecībā. Viduslaikos arhitekti centās panākt ideālas gotikas katedrāļu proporcijas, bet 20. gadsimta industrializācijas laikā tika izveidotas standartizētu normu un noteikumu sistēmas, kas kļuva par pamatu masveida būvniecībai

Mūsdienās attīstības vektors mainās no striktas cēloņsakarību meklēšanas uz varbūtības analīzi, korelāciju un slēpto likumsakarību meklēšanu. Nozarē sākas jauns posms - dati kļūst par galveno resursu, un uz tiem balstīta analīze aizstāj intuitīvās un lokāli optimizētās pieejas.



Attēls 10.1-1 Būvniecības datu slēptais potenciāls: uzņēmumā esošie aprēķini ir tikai aisberga virsotne, ko vadība var analizēt.

Uzņēmuma informācijas sistēma ir kā aisbergs ( 10.1-1. attēls): tikai neliela datu potenciāla daļa ir redzama uzņēmuma vadībai, bet galvenā vērtība slēpjас dzījumos. Ir svarīgi novērtēt datus ne tikai pēc to pašreizējās izmantošanas, bet arī pēc iespējām, ko tie atklās nākotnē. Tieši tie uzņēmumi, kas iemācīsies iegūt slēptos modeļus un radīt jaunas atziņas no datiem, radīs ilgtspējīgas konkurences priekšrocības

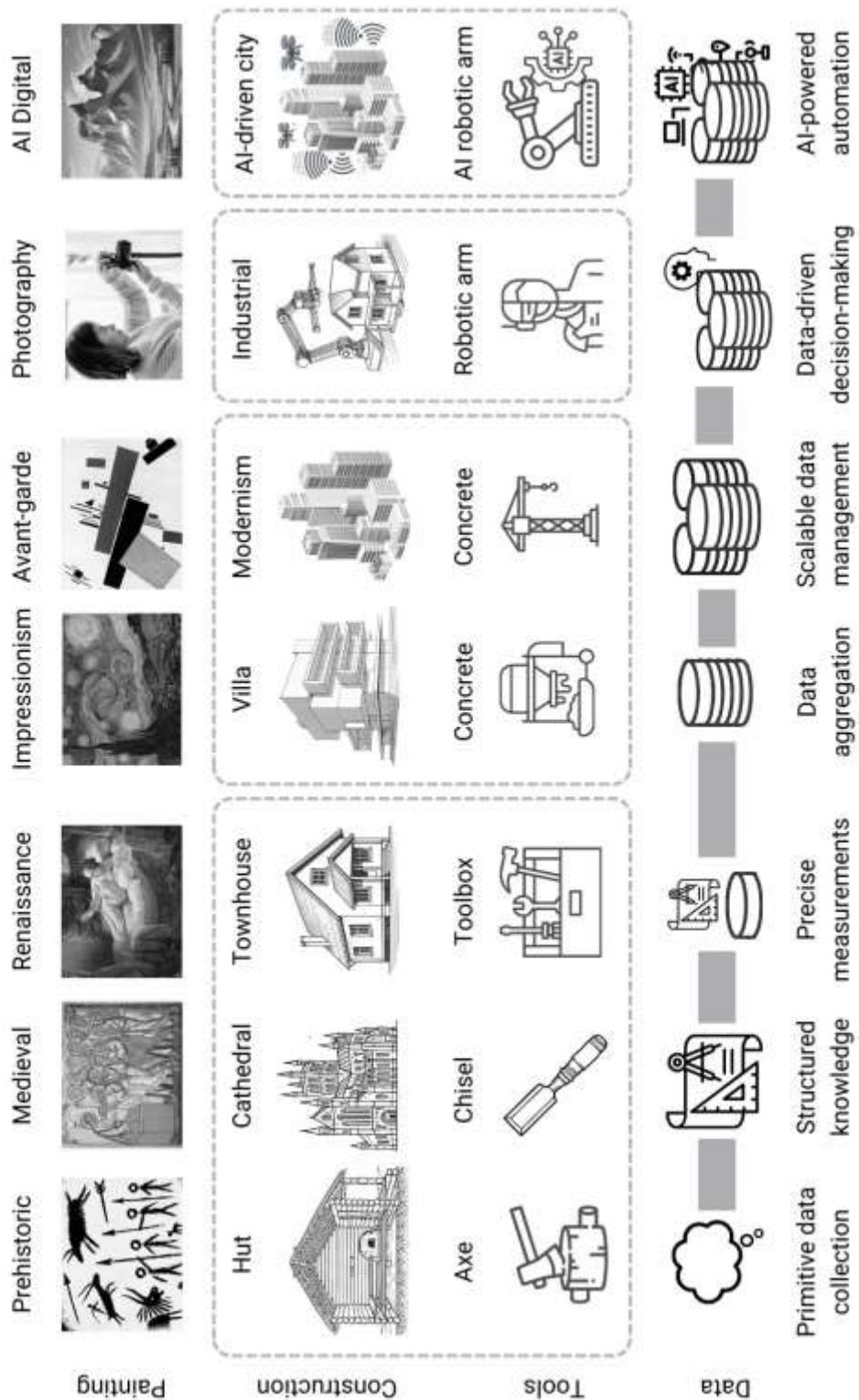
Slēpto likumsakarību meklēšana un datu jēgas veidošana nav tikai darbs ar skaitļiem, bet gan radošs process, kas prasa abstraktu domāšanu un spēju saskatīt kopainu aiz atšķirīgiem elementiem. Šajā ziņā darba ar datiem evolūciju var salīdzināt ar vizuālās mākslas evolūciju ( 10.1-2. attēls).

Būvniecības attīstība ir ļoti līdzīga tēlotājmākslas attīstībai. Abos gadījumos cilvēce ir attīstījusies no primitīvām metodēm līdz sarežģītām vizualizācijas un analīzes tehnoloģijām. Aizvēsturiskajos laikos cilvēki ikdienas uzdevumu risināšanai izmantoja alu zīmējumus un primitīvus instrumentus. Viduslaikos un renesances laikā arhitektūras un mākslas izsmalcinātības līmenis krasī pieauga. Līdz agrajiem viduslaikiem būvniecības rīki bija attīstījušies no vienkārša cirvja līdz plašam instrumentu komplektam, kas simbolizēja tehnisko zināšanu izaugsmi.

Reālisma laikmets bija pirmsais apvērsums vizuālajā mākslā: mākslinieki iemācījās atveidot vissīkākās detaļas, panākot maksimālu ticamību. Būvniecībā šī perioda analogs bija precīzas inženiertehniskās metodes, detalizēti rasējumi un stingri reglamentēti aprēķini, kas gadsimtiem ilgi kļuva par projektēšanas prakses pamatu.

Vēlāk impresionisms mainīja mākslinieciskās realitātes uztveri: tā vietā, lai burtiski atveidotu formu, mākslinieki sāka attēlot noskaņu, gaismu un dinamiku, cenšoties atspoguļot kopējo iespaidu, nevis absolūtu precizitāti. Līdzīgi arī mašīnmācīšanās ēku analīzē virzās no stingriem loģiskiem modeļiem uz rakstu atpazīšanu un varbūtības modeļiem, kas jauj "ieraudzīt" datos apslēptas atkarības, kas klasiskajā analīzē nav pieejamas. Šī pieeja atbalso Bauhaus minimālisma un funkcionalitātes idejas, kur nozīme (funkcija) ir svarīgāka par formu. Bauhauss centās likvidēt lieko, atteikties no ornamentiem skaidrības, utilitarisma un masu pievilcības labad. Lietām bija jābūt saprotamām un noderīgām, bez liekā - estētika radās no dizaina loģikas un mērķa.

Līdz ar fotografijas parādīšanos 19. gadsimta beigās māksla ieguva jaunu rīku, ar kura palīdzību varēja iemūžināt realitāti ar vēl nebijušu precizitāti, un apvērsa attieksmi pret vizuālo mākslu. Līdzīgi arī būvniecībā 21. gadsimta industriālā revolūcija novēd pie robotizētu tehnoloģiju, lāzeru, IoT, RFID un tādu koncepciju kā Connected Construction (savienotā būvniecība) izmantošanas, kur atsevišķu parametru apkopošana ir attīstījusies līdz mērogojamai intelīgentai pilnas būvlaukuma realitātes fiksēšanai.



Attēls 10.1-2 Tēlotājmākslas evolūcijas ēra saskan ar to, kā attīstās pieeja darbam ar datiem būvniecības nozarē.

Šodien, līdzīgi kā vizuālajā mākslā notiek pārvērtības, parādoties mākslīgā intelekta rīkiem un LLM, būvniecības nozare piedzīvo vēl vienu kvantu lēcienu: inteliģentas sistēmas, ko vada mākslīgais intelekts (AI), LLM tērzēšana ļauj prognozēt, optimizēt un ģenerēt risinājumus ar minimālu cilvēka iejaukšanos.

Datu loma projektēšanā un pārvaldībā ir radikāli mainījusies. Ja agrāk zināšanas tika nodotas mutiski un empīriski - līdzīgi kā līdz 19. gadsimtam realitāte tika fiksēta ar rokām gleznotās bildēs -, tad mūsdienās galvenā uzmanība tiek pievērsta pilnīgai būvniecības "attēla" digitalizācijai. Ar mašīnmācīšanās algoritmu palīdzību šī digitālā bilde tiek pārveidota par būvniecības realitātes impresionistisku attēlojumu - nevis precīzu repliku, bet gan vispārinātu, varbūtisku izpratni par procesiem.

Mēs strauji tuvojamies laikmetam, kad ēku projektēšanas, būvniecības un ekspluatācijas procesus ne tikai papildinās, bet lielā mērā vadīs mākslīgā intelekta sistēmas. Līdzīgi kā mūsdienu digitālā māksla tiek radīta bez otas - izmantojot teksta norādes un ģeneratīvus modeļus - nākotnes arhitektūras un inženiertehniskos risinājumus veidos lietotāja noteikti galvenie pieprasījumi un parametri.

21. gadsimtā piekļuve datiem, to interpretācija un analītikas kvalitāte kļūst neaizstājama, lai projekts būtu veiksmīgs. Un datu vērtību nosaka nevis to apjoms, bet gan speciālistu spēja tos analizēt, pārbaudīt un pārvērst darbībā.

### **Uz datiem balstīta pieeja būvniecībā: jauns infrastruktūras līmenis**

Cilvēces vēsturē katrs šāds tehnoloģiskais lēciens ir būtiski mainījis ekonomiku un sabiedrību. Šodien mēs esam liecinieki jaunam pārmaiņu vilnim, kas pēc mēroga ir salīdzināms ar 19. gadsimta industriālo revolūciju. Tomēr, ja pirms simts gadiem galvenais pārmaiņu virzītājspēks bija mehāniskie spēki un energotehnoloģijas, tad tagad tie ir dati un mākslīgais intelekts.

Mašīnmācīšanās, LLM un mākslīgā intelekta aģenti maina pašu lietojumprogrammu būtību, padarot nevajadzīgus tradicionālos programmatūras kopumus (kas aplūkoti grāmatas otrajā daļā) (2.2-3. att.). Visa datu loģika ir centrēta mākslīgā intelekta aģentos, nevis grūti kodētos uzņēmējdarbības noteikumos (2.2-4. attēls).

Datu laikmetā tradicionālais skatījums uz lietojumprogrammām tiek būtiski pārveidots. Mēs virzāmies uz modeli, kurā apjomīgas, modulāras uzņēmumu sistēmas neizbēgami aizstās vietu atvērtiem, viegliem, pielāgotiem risinājumiem.

Nākotnē saglabāsies tikai pamatā esošā datu struktūra, un visa mijiedarbība ar to notiks, izmantojot aģentus, kas strādās tieši ar datubāzi. Es patiešām ticiu, ka viss lietojumprogrammu kaudze izzudīs, jo tā vienkārši nebūs vajadzīga, kad mākslīgais intelekts mijiedarbībās tieši ar pamatā esošo datubāzi. Visu savu karjeru esmu pavadījis, strādājot SaaS - veidojot uzņēmumus, strādājot tajos, un, godīgi sakot, es, iespējams, šobrīd neuzsāktu jaunu SaaS uzņēmumu. Un es, iespējams, šobrīd arī neinvestētu SaaS uzņēmumos. Situācija ir pārāk neskaidra. Tas nenozīmē, ka nākotnē nebūs programmatūras uzņēmumu, tikai tie izskatīsies pavisam citādi. Nākotnes sistēmas būs datubāzes ar biznesa loģiku, kas iekļauta [AI] aģentos. Šie aģenti vienlaikus strādās ar vairākām datu krātuvēm, neaprobežojoties tikai ar vienu datu bāzi. Visa loģika tiks pārcelta uz mākslīgā intelekta slāni [46].

- Metjū Bermans, izpilddirektors Forward Future

Jaunās paradigmas galvenā atšķirība ir tehnoloģiskā balasta samazināšana līdz minimumam. Monumentālu sarežģītu un slēgtu programmatūras sistēmu vietā mēs iegūsim elastīgus, atvērtus un ātri pielāgojamus moduļus, kas burtiski "dzīvo" datu plūsmā ( 7.4-1. attēls - Apache Airflow, NiFi). Nākotnes procesu pārvaldības arhitektūra paredz izmantot mikroprogrammas - kompaktus, mērķtiecīgus rīkus, kas būtiski atšķiras no masveida un slēgtām ERP, PMIS, CDE, CAFM sistēmām. Jaunie aģenti būs pēc iespējas adaptīvāki, integrēti un pielāgoti konkrētam uzņēmumam (piemēram, Low-Code/No-Code 7.4-6. attēls).

*Visa biznesa loģika tiks nodota šiem [AI] aģentiem, un šie aģenti veiks CRUD [Create, Read, Update, and Delete] operācijas vairākos repositorijos, t. i., tie neatšķirs, kurš backend tiek izmantots. Tie atjauninās vairākas datubāzes, un visa loģika nonāks tā sauktajā AI -līmenī. Un, kad visa loģika atradīsies AI līmenī, cilvēki sāks nomainīt backendus. Mēs jau tagad novērojam diezgan lielu tirgus uzvaras procentuālo daļu Dynamics backends un aģentu izmantošanā, un mēs agresīvi virzīsimies šajā virzienā, cenšoties to visu apvienot. Gan klientu apkalpošanā, gan citās jomās, piemēram, ne tikai CRM, bet arī mūsu finanšu un operāciju risinājumos. Jo cilvēki vēlas vairāk uz mākslīgo intelektu balstītu biznesa lietojumprogrammu, kurās loģisko slāni var vadīt mākslīgais intelekts un mākslīgā intelekta aģenti. [...]. Viena no aizraujošākajām lietām man ir Excel ar Python, kas ir salīdzināma ar GitHub ar Copilot. Es gribu teikt, ka mēs esam izdarījuši tā: tagad, kad jums ir Excel, vienkārši atveriet to, palaidiet Copilot un sāciet ar to spēlēties. Tā vairs nav tikai pieejamo skaitļu izpratne - tā pati izveidos plānu. Tāpat kā GitHub Copilot darbvieta izveido plānu un pēc tam to izpilda, tā ir kā datu analītikis, kas izmanto Excel kā rīku, lai vizualizētu rindas un kolonnas analīzei. Tātad Copilot izmanto Excel kā rīku ar visām tā iespējām, jo tas var ġenerēt datus un tam ir Python interpretētājs.*

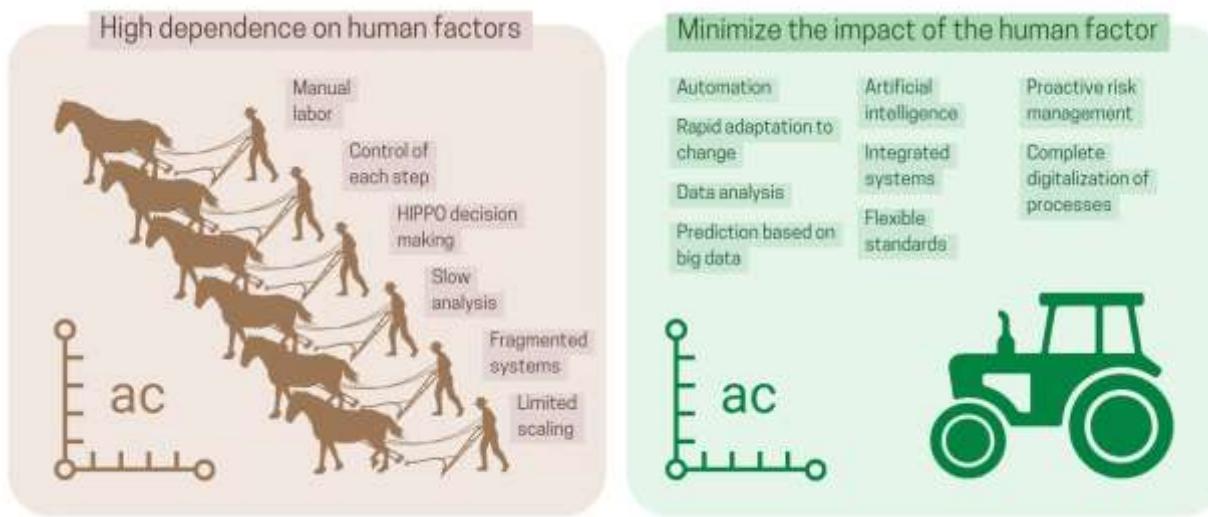
- Satya Nadella, izpilddirektors, Microsoft, intervija ar BG2 kanālu, 2024. gada decembris.  
[28]

Biroja lietojumprogrammu loģikas pārveide - pāreja no modulārām, slēgtām sistēmām uz mākslīgā intelekta aģentiem, kas strādā tieši ar atvērtiem datiem, - ir tikai daļa no daudz plašāka procesa. Runa nav tikai par saskarņu vai programmatūras arhitektūras maiņu - pārmaiņas ietekmēs darba

organizācijas, lēmumu pieņemšanas un uzņēmējdarbības vadības pamatprincipus. Būvniecībā tas novedīs pie uz datiem balstītas loģikas, kurā dati kļūs par centrālo elementu procesos no projektēšanas līdz resursu pārvaldībai un būvniecības uzraudzībai.

## Nākamās paaudzes digitālais birojs: kā mākslīgais intelekts maina darba vidi

Gandrīz pirms gadsimta cilvēce jau piedzīvoja līdzīgu tehnoloģisko revolūciju. Pāreja no tvaika dzinējiem uz elektromotoriem aizņēma vairāk nekā četras desmitgades, bet galu galā veicināja vēl nebijušu produktivitātes pieaugumu - galvenokārt pateicoties energijas jaudas decentralizācijai un jauno risinājumu elastībai. Šī pāreja ne tikai mainīja vēstures gaitu, pārvietojot lielāko iedzīvotāju daļu no lauku apvidiem uz pilsētām, bet arī lika pamatus mūsdienu ekonomikai. Tehnoloģiju vēsture ir ceļojums no fiziskā darba līdz automatizācijai un viedām sistēmām. Tāpat kā traktors nomainīja daudzus augsnēs apstrādātājus, modernās digitālās tehnoloģijas aizstāj tradicionālās biroju būvniecības vadības metodes ( 10.1-3. attēls). Vēl 20. gadsimta sākumā lielākā daļa pasaules iedzīvotāju zemi apstrādāja ar rokām, līdz 20. gadsimta 30. gados sākās darba mehanizācija ar mašīnām un traktoriem.



**Attēls 10.1-3 Tāpat kā 20. gadsimta sākumā traktors aizstāja desmitiem cilvēku, 21. gadsimtā mašīnmācīšanās aizstās tradicionālās biznesa un projektu vadības metodes.**

Tāpat kā pirms simts gadiem cilvēce no individuālu zemes gabalu apstrādes ar primitīviem darbarīkiem pārgāja uz liela mēroga lauksaimniecību, izmantojot tehniku, šodien mēs pārejam no atšķirīgu informācijas "siloso" apstrādes uz darbu ar datu masīviem, izmantojot jaudīgus "traktorus" - ETL cauruļvadu un mākslīgā intelekta algoritmus.

Mēs atrodamies uz līdzīga lēciena sliekšņa, taču digitālā līmenī: no tradicionālās, manuālās uzņēmējdarbības pārvaldības uz datu balstītiem modeļiem.

Lai nonāktu līdz pilnvērtīgai uz datiem balstītai arhitektūrai, būs nepieciešams laiks, ieguldījumi un organizatoriski pūliņi. Taču šis ceļš paver ceļu ne tikai pakāpeniskai uzlabošanai, bet arī kvalitatīvam lēcienam uz būvniecības procesu lielāku efektivitāti, pārredzamību un vadāmību. Tas viss ir atkarīgs

no sistemātiskas digitālo rīku ieviešanas un atteikšanās no novocojušās uzņēmējdarbības prakses.

Uzdevumu parametrizācija, ETL, LLM, IoT komponenti, RFID, tokenizācija, lielie **dati** un mašīnmācīšanās pārveidos tradicionālo būvniecību par uz **datiem balstītu būvniecību**, kurā katru projekta un būvniecības uzņēmējdarbības detaļu kontrolēs un optimizēs dati.

Agrāk informācijas analīzei bija nepieciešamas tūkstošiem cilvēkstundu. Tagad šos uzdevumus veic algoritmi un LLM, kas, izmantojot uzvednes, pārvērš atšķirīgas datu kopas stratēģiskos avotos. Tehnoloģiju pasaulē notiek tas pats, kas notika laukaimniecībā: mēs pārejam no kapļiem uz automatizētu agrobūvniecību. Tāpat arī biroja darbs būvniecībā - no Excel failiem un manuālas apkopošanas - pāriet uz inteliģentu sistēmu, kurā dati tiek apkopoti, attīrti, strukturēti un pārvērsti atziņās.

Jau šodien uzņēmumiem būtu jāsāk "kultivēt" informācijas laukus, veicot kvalitatīvu datu vākšanu un informācijas strukturēšanu, un "mēslot" tos ar tīrišanas un normalizēšanas rīkiem, un pēc tam "novākt" to ražu, izmantojot prognozēšanas analītiku un automatizētus risinājumus. Ja mūsdienīgs laukaimnieks ar mašīnu spēj aizstāt simts augstes apstrādātājus, tad inteliģenti algoritmi spēs noņemt rutīnu no darbiniekiem un nodot viņiem informācijas plūsmu stratēģisko pārvaldnieku lomu.

Tomēr ir svarīgi saprast, ka patiesi uz datiem balstītas organizācijas izveide nav ātrs process. Tas ir ilgtermiņa stratēģisks virziens, līdzīgs jaunas vietas izveidei jauna meža ( 1.2-5. attēls) sistēmu audzēšanai, kur katrs "koks" šajā ekosistēmā ir atsevišķs process, kompetence vai rīks, kura izaugsmei un attīstībai nepieciešams laiks. Tāpat kā ūsta meža gadījumā, panākumi ir atkarīgi ne tikai no stādāmā materiāla kvalitātes (tehnoloģijas), bet arī no augstnes (uzņēmuma kultūra), klimata (uzņēmējdarbības vide) un kopšanas (sistēmiskā pieeja).

Uzņēmumi vairs nevarēs paļauties tikai uz slēgtiem, "out-of-the-box" risinājumiem. Atšķirībā no iepriekšējiem tehnoloģiju attīstības posmiem pašreizējo pāreju uz atvērtu piekļuvi datiem, mākslīgo intelektu un atklātā pirkoda programmatūru, visticamāk, neatbalstīs lielie ražotāji, jo tā tieši apdraud to iedibinātos uzņēmējdarbības modeļus un galvenos ieņēmumu avotus.

Kā liecina Hārvarda Biznesa skolas pētījums [40], kas jau tika aplūkots nodaļā par ceturto un piekto tehnoloģisko revolūciju, izmaksas visu uzņēmumu visvairāk izmantoto atvērtā koda risinājumu izveidei no nulles būtu aptuveni 4,15 miljardi dolāru. Taču, ja mēs iedomājamies, ka katrs uzņēmums pats izstrādātu savas alternatīvas bez piekļuves esošajiem Open Source rīkiem, kā tas ir noticed pēdējās desmitgadēs, kopējās uzņēmējdarbības izmaksas varētu sasniegt kolosālos 8,8 triljonus dolāru - tā ir iracionālā pieprasījuma cena, par kādu var novērtēt programmatūras tirgu.

Tehnoloģiskais progress neizbēgami novēdīs pie iedibināto uzņēmējdarbības modeļu pārskatīšanas. Ja agrāk uzņēmumi varēja pelnīt naudu, izmantojot sarežģītus, nepārskatāmus procesus un slēgtus datus, tad, attīstoties mākslīgajam intelektam un analītikai, šāda pieeja kļūst arvien mazāk dzīvotspējīga.

Piekļuves datiem un rīkiem demokratizācijas rezultātā tradicionālais programmatūras pārdošanas tirgus var ievērojami sarukt. Tomēr tajā pašā laikā pieauga jauns tirgus - digitālo zināšanu, pielāgošanas, integrācijas un risinājumu izstrādes tirgus. Šeit vērtību radīs nevis licenču pārdošana, bet gan spēja veidot elastīgus, atvērtus un pielāgojamus digitālos procesus. Tāpat kā elektrifikācija un traktoru parādīšanās ir radījusi jaunas nozares, tāpat arī lielo datu, mākslīgā intelekta un LLM izmantošana pavērs pilnīgi jaunus uzņēmējdarbības horizontus būvniecības nozarē, kas prasīs ne tikai tehnoloģiskus ieguldījumus, bet arī dziļu domāšanas veida, procesu un organizatorisko struktūru transformāciju. Un tie uzņēmumi un profesionāļi, kas to apzinās un sāks rīkoties jau šodien, būs rītdienas līderi.

Pasaulē, kurā atvērtie dati klūst par galveno vērtību, informācijas pieejamība mainīs spēles noteikumus. Investori, klienti un regulatori arvien vairāk pieprasīs pārredzamību, un mašīnmācīšanās algoritmi spēs automātiski noteikt neatbilstības aplēsēs, termiņos un izmaksās. Tas rada priekšnoteikumus jaunam digitālās transformācijas posmam, kas pakāpeniski noved mūs pie būvniecības nozares "uberizācijas".

## **Atvērtie dati un Uberizācija ir drauds esošajiem būvniecības uzņēmumiem**

Būvniecība klūst par informācijas pārvaldības procesu. Jo precīzāki, kvalitatīvāki un pilnīgāki ir dati, jo efektīvāka ir ēku projektēšana, aprēķini, izmaksu aprēķini, būvniecība un ekspluatācija. Nākotnē galvenais resurss nebūs celtnis, bet gan betons un armatūra, bet gan spēja apkopot, analizēt un izmantot informāciju.

Nākotnē būvniecības uzņēmumu klienti - investori un klienti, kas finansē būvniecību - neizbēgami izmantos atvērto datu un analītikas vēsturisko datu vērtību. Tas pavērs iespējas automatizēt projektu laika grafiku un izmaksu aprēķinus, neiesaistot būvniecības uzņēmumus izmaksu aprēķināšanas jautājumos, kas palīdzēs kontrolēt izmaksas un ātrāk identificēt liekās izmaksas.

Iedomājieties būvlaukumu, kurā lāzerskeneri, kvadrokopteri un fotogrammetrijas sistēmas apkopo precīzus reāllaika datus par izmantotā betona daudzumu. Šī informācija tiek automātiski pārvērsta vienkāršos plakanos MESH modeļos ar metadatiem, apejot apgrūtinošās CAD (BIM) sistēmas, bez atkarības no sarežģītiem ģeometriskiem kodoliem, ERP vai PMIS. Šie būvlaukumā savāktie dati tiek centralizēti pārsūtīti uz vienotu strukturētu repozitoriju, kas ir pieejams klientam neatkarīgai analīzei, kur tiek augšupielādētas reālās cenas no dažādiem būvdarbu veikalim ar dažādiem parametriem, sākot no kreditu finansēšanas likmes līdz dinamiski mainīgiem faktoriem, piemēram, laikapstākļiem, būvmateriālu biržas kotācijām, logistikas tarifiem un statistiskajām sezonālajām darbaspēka cenu svārstībām. Šādos apstākļos jebkādas neatbilstības starp projektēto un faktisko materiālu apjomu kļūst uzreiz acīmredzamas, padarot neiespējamu manipulēt ar tāmēm gan projektēšanas posmā, gan projekta nodošanas brīdī. Rezultātā būvniecības procesa pārredzamību nodrošina nevis uzraugu un vadītāju armija, bet gan objektīvi digitālie dati, kas līdz minimumam samazina cilvēcisko faktoru un

spekulāciju iespēju.

Nākotnē šāda veida datu kontroles darbu drīzāk veiks datu pārvaldnieki klienta pusē ( 1.2-4. attēls CQMS pārvaldnieks). Īpaši tas attiecas uz aprēķiniem un projektu tāmēm: tur, kur agrāk bija vesels tāmētāju departaments, jau rīt būs mašīnmācīšanās un prognozēšanas rīki, kas noteiks cenu robežas, kurās celtniecības uzņēmumi varēs iekļauties.

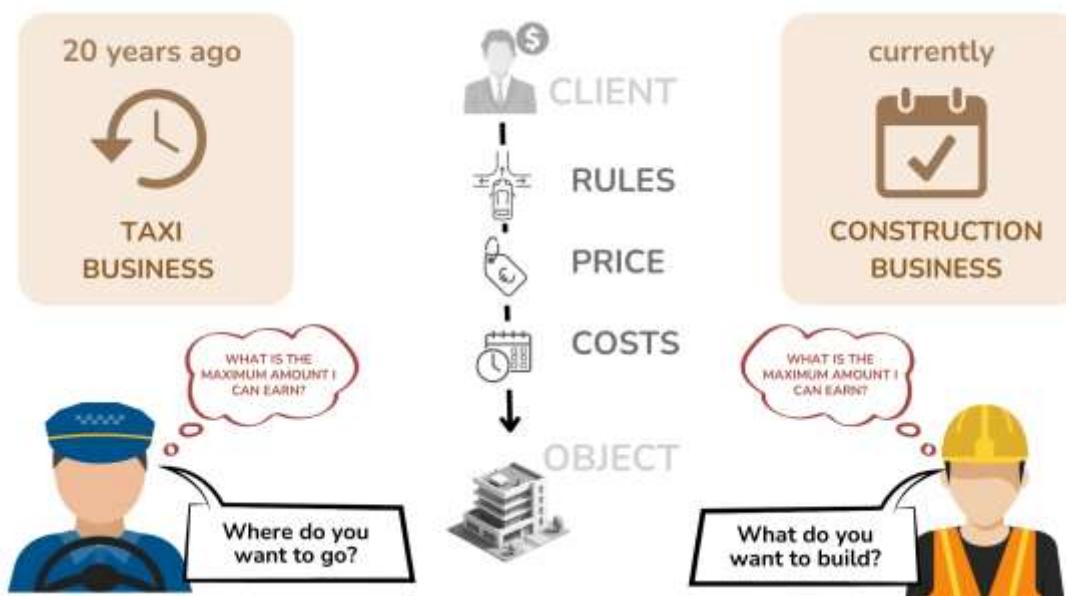
*Nemot vērā [būvniecības] nozares sadrumstalotību, kur lielāko daļu sistēmu un apakšsistēmu piegādā MVU, digitālā stratēģija ir jāveido no klienta. Klientiem ir jārada apstākļi un mehānismi, lai atraisītu piegādes kēdes digitālās spējas [20].*

- Endrjū Deiviss un Džuliāno Deničols, Accenture "Lielākas vērtības radīšana, izmantojot kapitālprojektus"

Šāda datu atklātība un pārredzamība apdraud būvniecības uzņēmumus, kuri ir pieraduši pelnīt naudu no nepārskatāmiem procesiem un neskaidriem pārskatiem, kur spekulācijas un slēptās izmaksas var slēpties aiz sarežģītiem un slēgtiem formātiem un modulārām patentētām datu platformām. Tāpēc būvniecības uzņēmumi, tāpat kā pārdevēji, kas popularizē atvērtā koda risinājumus, visticamāk, nebūs ieinteresēti pilnībā ieviest atvērtos datus savos uzņēmējdarbības procesos. Ja dati ir pieejami un viegli apstrādājami klientam, tos var pārbaudīt automātiski, tādējādi novēršot iespēju pārvērtēt apjomus un manipulēt ar tāmēm.

Saskaņā ar Pasaules Ekonomikas foruma ziņojumu "Veidojot būvniecības nākotni" (2016) [5], viens no galvenajiem nozares izaicinājumiem joprojām ir pasīvā klienta loma. Tomēr tieši pasūtītājiem būtu jāuzņemas lielāka atbildība par projektu iznākumu - sākot ar agrīnu plānošanu, ilgtspējīgu mijiedarbības modeļu izvēli un beidzot ar izpildes uzraudzību. Bez projektu īpašnieku aktīvas līdzdalības nav iespējama būvniecības nozares sistēmiska pārveidošana.

Apjoma un izmaksu aprēķinu kontroles zaudēšana pēdējo 20 gadu laikā jau ir pārveidojusi citas nozares, ļaujot klientiem tieši, bez starpniekiem, stimulēt savus mērķus. Digitalizācija un datu pārredzamība ir pārveidojusi daudzus tradicionālos uzņēmējdarbības modeļus, piemēram, taksometru vadītājus līdz ar Uber parādīšanos ( 10.1-4. attēls), viesnīcu īpašniekus līdz ar Airbnb parādīšanos un mazumtirgotājus un veikalus līdz ar Amazon parādīšanos, kā arī bankas līdz ar neo banku un decentralizētu fintech ekosistēmu parādīšanos, kur tieša piekļuve informācijai un laika un izmaksu aprēķinu automatizācija ir ievērojami samazinājusi starpnieku lomu.



Attēlā: 10.1-4 Būvniecības bizness saskarsies ar überizāciju, ar kādu pirms 10 gadiem nācās saskarties taksometru vadītājiem, viesnīcu īpašniekiem un mazumtirgotājiem.

Piekļuves datiem un to apstrādes rīkiem demokratizācijas process ir neizbēgams, un laika gaitā atvērtie dati par visiem projekta komponentiem kļūs par klientu prasību un jaunu standartu. Tāpēc jautājumus par atvērtu formātu un pārredzamu aprēķinu ieviešanu veicinās investori, klienti, bankas un privātā kapitāla fondi (privātā kapitāla fondi) - tie, kas ir izbūvēto objektu galalietotāji un pēc tam desmitiem gadu ekspluatē objektu.

Lielākie investori, klienti un bankas jau pieprasā būvniecības nozares pārredzamību. Saskaņā ar Accenture pētījumu "Radīt lielāku vērtību, izmantojot kapitālprojektus" (2020) [20], pārredzami un uzticami dati kļūst par izšķirošu faktoru investīciju lēmumu pieņemšanā būvniecības nozarē. Kā norāda eksperti, uzticama un efektīva projektu vadība nav iespējama bez pārredzamības, jo īpaši krīzes situācijās. Turklat aktīvu īpašnieki un būvuzņēmēji aizvien vairāk virzās uz līgumiem, kas stimulē datu koplietošanu un sadarbības analīzi, atspoguļojot pieaugošās prasības no investoru, banku un regulatoru puses pēc atbildības un pārredzamības.

Investora, pasūtītāja pārvietošanās no idejas līdz pabeigtai ēkai nākotnē būs līdzīga ceļošanai ar autopilotu palīdzību - bez vadītāja būvfirma formā, sola kļūt neatkarīga no spekulācijām un nenoteiktības.

Atvērto datu un automatizācijas ēra neizbēgami mainīs būvniecības nozari, līdzīgi kā tas jau ir noticis banku, tirdzniecības, lauksaimniecības un loģistikas nozarē. Šajās nozarēs starpnieku loma un tradicionālie uzņēmējdarbības veidi piekāpjas automatizācijai un robotizācijai, neatstājot vietu nepamatotiem uzcenojumiem un spekulācijām.

Dati un procesi visās cilvēku saimnieciskajās darbībās ne ar ko neatšķiras no tā, ar ko nākas saskarties

būvniecības nozares profesionāļiem. Ilgtermiņā būvniecības uzņēmumi, kas šodien dominē tirgū, nosakot cenas un pakalpojumu kvalitātes standartus, var zaudēt savu galveno starpnieka lomu starp klientu un viņa būvniecības projektu.

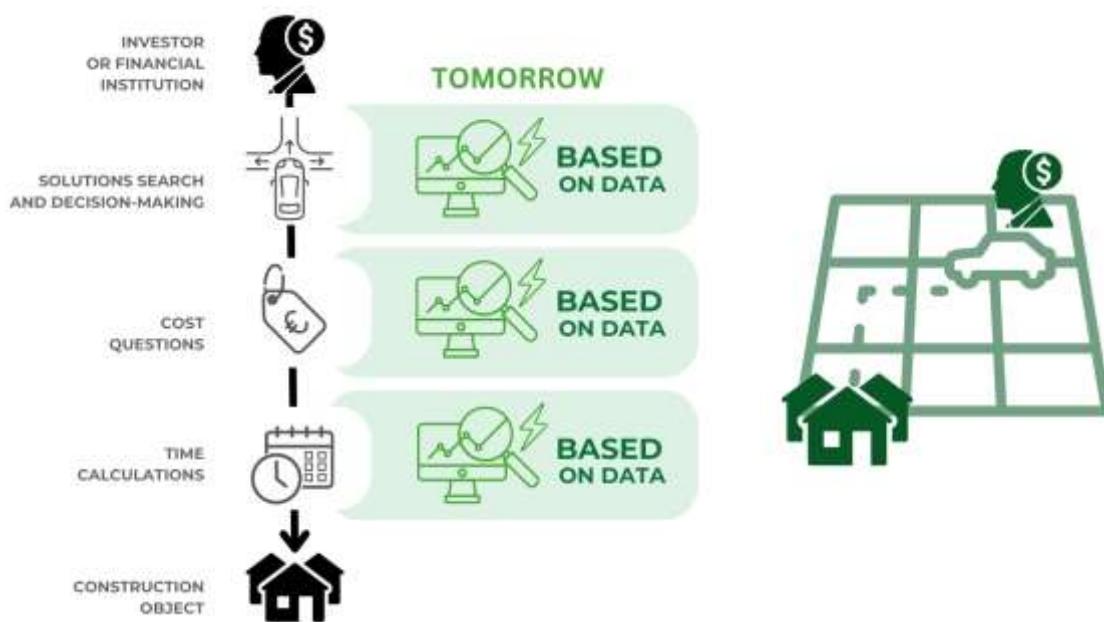
## **Neatrisinātās überizācijas problēmas kā pēdējā iespēja izmantot laiku pārveidošanai**

Taču atgriezīsimies pie būvniecības nozares realitātes. Lai gan dažās ekonomikas nozarēs parādās pašbraucošas automašīnas, decentralizētas finanšu sistēmas un uz mākslīgo intelektu balstīti risinājumi, ievērojama daļa būvniecības uzņēmumu joprojām ir organizācijas, kas darbojas papīra formātā un kurās galvenie lēmumi tiek pieņemti, vairāk balstoties uz atsevišķu speciālistu intuīciju un pieredzi.

Šajā paradigmā mūsdienu būvniecības uzņēmumu var salīdzināt ar 20 gadus vecu taksometru uzņēmumu, kas kontrolē resursus, maršrutus un piegādes laikus un ir atbildīgs par "brauciena" laiku un izmaksām - no projekta idejas (loģistikas un uzstādīšanas procesa) līdz projekta piegādei. Tāpat kā GPS (būvniecībā IoT, RFID) un mašīnmācīšanās algoritmi laika/izmaksu aprēķinos savulaik pārveidoja transporta jomu, datiem, algoritmiem un mākslīgā intelekta -aģentiem ir potenciāls pārveidot būvniecības pārvaldību - no intuitīviem novērtējumiem līdz prognozējošiem, vadītiem modeļiem. Pēdējo 20 gadu laikā daudzās nozarēs - finanšu, lauksaimniecības, mazumtirdzniecības un loģistikas - ir pakāpeniski izzudusi spēja spekulēt, izmantojot datu nepārredzamību. Cenas, piegādes izmaksas vai finanšu darījumi tiek aprēķināti automātiski un statistiski pamatooti - dažās sekundēs digitālajās platformās.

Raugoties nākotnē, būvniecības uzņēmumiem jāapzinās, ka, demokratizējot piekļuvi datiem un to analīzes rīkiem, tiks izjaukta tradicionālā pieeja projektu izmaksu un termiņu aplēsēm un novērsta iespēja spekulēt ar nepārredzamiem datiem par apjomu un cenām.

Tāpat kā braukšana pa regulētu ceļu bez autovadītāja iejaukšanās, arī būvniecības procesi nākotnē arvien vairāk līdzināsies "Uber" sistēmai - ar automatizētu laika un izmaksu aprēķinu, pārredzamu uzdevumu maršrutēšanu un minimālu atkarību no cilvēciskā faktora. Tas mainīs "brauciena" no idejas līdz realizācijai būtību, padarot to paredzamāku, pārvaldāmāku un uz datiem balstītu.



**Attēls 10.1-5 Izmaksas un "ceļošanas" laiks būvniecības laikā tiks noteikts, izmantojot mašīnmācīšanās un statistikas rīkus.**

Līdz ar pakāpenisku jaunu noteikumu un prasību ieviešanu gandrīz visās pasaules valstīs, kas paredz CAD- (BIM)-modeļu nodošanu klientiem vai bankām, kas finansē būvniecības projektus, klientam un pasūtītājam ir iespēja patstāvīgi nodrošināt izmaksu un apjoma aprēķinu pārredzamību. Tas ir īpaši svarīgi lielajiem pasūtītājiem un investoriem, kuriem ir pietiekama kompetence un rīki, lai operatīvi analizētu apjomus un uzraudzītu tirgus cenas. Uzņēmumiem, kas īsteno liela mēroga standarta projektus - veikalus, biroju ēkas, dzīvojamo ēku kompleksus -, šāda prakse kļūst par standartu.

Tā kā modeļu informācijas saturs kļūst pilnīgāks un standartizētāks, manipulāciju un spekulāciju iespēja praktiski izzūd. Digitālā transformācija pakāpeniski maina spēles noteikumus būvniecības nozarē, un uzņēmumi, kas nepielāgojas šīm pārmaiņām, var saskarties ar nopietnām problēmām.

Pieaugošā konkurence, tehnoloģiju radīti traucējumi un peļņas samazināšanās var ietekmēt uzņēmējdarbības ilgtspēju. Nemot vērā ierobežoto likviditāti, arvien vairāk nozares dalībnieku pievērsas automatizācijai, analītikai un datu tehnoloģijām, lai uzlabotu efektivitāti un procesu pārredzamību. Šie rīki kļūst par svarīgu resursu, lai saglabātu konkurētspēju mainīgajā ekonomiskajā vidē.

Iespējams, mums nevajadzētu gaidīt, kad ārējie apstākļi piespiedīs steidzami rīkoties - daudz efektīvāk ir sākt gatavoties jau šodien, stiprinot digitālās kompetences, ieviešot mūsdienīgus risinājumus un veidojot uz datiem orientētu kultūru.

Viens no pēdējiem būtiskākajiem tehnoloģiskajiem šķēršļiem būvniecības nozares plaša mēroga digitālajai transformācijai, kas tuvākajos gados skars ikvienu uzņēmumu, ir būvniecības projektu elementu automātiskās klasifikācijas problēma.

Bez uzticamas, precīzas un mērogojamas klasifikācijas nav iespējams radīt pamatu pilnvērtīgai analīzei, procesu automatizācijai un dzīves cikla pārvaldībai, izmantojot mākslīgo intelektu un prognozēšanas modeļus. Kamēr objektu klasifikācija joprojām ir atkarīga no manuālas interpretācijas, ko veic pieredzējuši speciālisti - meistari, projektētāji, tāmētāji -, būvniecības nozarei joprojām ir iespēja. Šo laiku var izmantot, lai sagatavotos neizbēgamajām pārmaiņām: pieaugošajām prasībām pēc pārredzamības, rīku un datu demokratizācijai un automātiskās klasifikācijas sistēmu parādīšanās, kas radikāli mainīs spēles noteikumus.

Būvniecības pasaules elementu automātiskās klasifikācijas uzdevums pēc sarežģītības ir salīdzināms ar objektu atpazīšanu bezpilota braukšanas sistēmās, kas ir viens no galvenajiem izaicinājumiem. Iedomāsimies bezpilota automobili, kas pārvietojas no punkta A uz punktu B ( 10.1-5. attēls). Pašreizējās automātiskās braukšanas sistēmas ir apgrūtinātas ar problēmu klasificēt objektus, kurus atpazīst lidaros un kamerās. Nepietiek ar to, ka automašīna vienkārši "redz" šķērsli vai orientieri, bet tai jāspēj atpazīt, vai tas ir gājējs, ceļa zīme vai atkritumu urna.

Ar līdzīgu būtisku izaicinājumu saskaras visa būvniecības nozare. Projekta elementus, piemēram, logus, durvis vai kolonas, var reģistrēt dokumentācijā, attēlot CAD modeļos, fotografēt būvlaukumā vai atpazīt lāzerskenēšanā iegūtos punktu mākoņos. Tomēr ar to vizuālu vai aptuvenu ģeometrisko atpazīšanu nepietiek, lai izveidotu patiesi automatizētu projektu vadības sistēmu. Ir nepieciešams nodrošināt, lai katrs elements tiktu precīzi un konsekventi iedalīts tipā, kas būs viennozīmīgi identificējams visos turpmākajos procesos - no tāmēm un specifikācijām līdz loģistikai, krājumu kontrolei un, kas ir vissvarīgākais, ekspluatācijai ( 4.2-6. attēls).

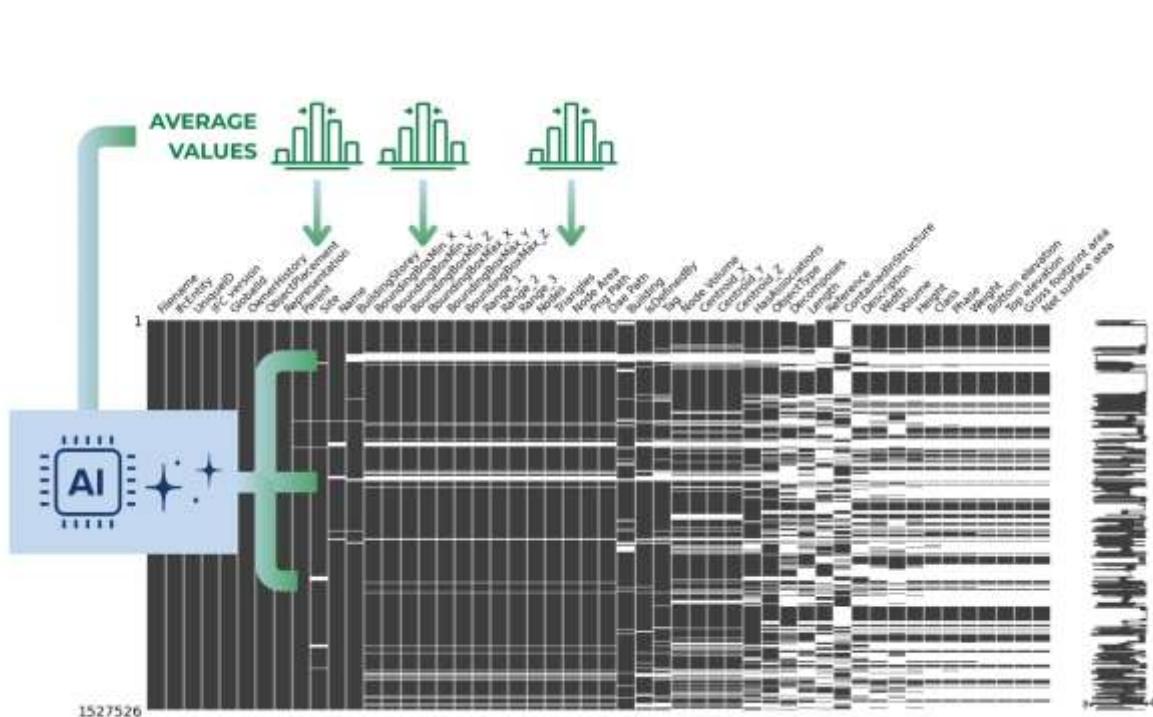
Tieši šajā posmā - pārejā no atzīšanas uz jēgpilnu klasifikāciju - rodas viens no galvenajiem šķēršļiem. Pat ja digitālās sistēmas ir tehniski spējīgas atšķirt un identificēt objektus modeļos un būvlaukumā, galvenās grūtības ir saistītas ar pareizu un kontekstuāli stabilu elementa tipa definēšanu dažādās programmatūras vidēs.. Piemēram, durvis projektētājs CAD modeļi var apzīmēt kā kategorijas "durvis" elementu, bet, pārnesot uz ERP vai PMIS sistēmu, tās var tikt nepareizi apzīmētas vai nu projektētāja klūdas dēļ, vai arī sistēmu nesakritību dēļ. Turklāt datu eksporta un importa laikā elements bieži zaudē dažus svarīgus atribūtus vai pazūd no sistēmas uzskaites. Tas rada nepilnības datu plūsmā un grauj būvniecības procesu digitalizācijas principu "no gala līdz galam". Tas rada kritisku plaisiru starp "redzamo" un "saprotamo" semantisko nozīmi, kas grauj datu integritāti un ievērojami sarežģī procesu automatizāciju visā būvniecības projekta dzīves ciklā.

Izaicinājuma risinājums universālai būvelementu klasifikācijai, izmantojot lielo datu un mašīnmācīšanās tehnoloģijas ( 10.1-6. attēls), būs nozares mēroga pārmaiņu katalizators - un, iespējams, negaidīts atklājums daudziem būvniecības uzņēmumiem. Vienota, mācāma klasifikācijas sistēma būs pamats mērogojamai analīzei, digitālajai pārvaldībai un mākslīgā intelekta ieviešanai būvniecības organizāciju ikdienas praksē.

NVIDIA un citi tehnoloģiju līderi jau piedāvā risinājumus citās nozarēs, kas spēj automātiski kategorizēt un strukturēt milzīgus teksta un vizuālās informācijas apjomus.

Piemēram, NVIDIA NeMo Curator modelis [161] specializējas automātiskā datu klasificēšanā un iedaļīšanā iepriekš noteiktās kategorijās, un tam ir būtiska nozīme informācijas apstrādes cauruļvadu optimizēšanā, lai precīzi pielāgotu un iepriekš apmācītu ģeneratīvos mākslīgā intelekta modeļus. Cosmos platforma tiek apmācīta, izmantojot reālās pasaules video un 3D ainas [162], nodrošinot pamatu autonomām sistēmām un digitālajiem dvīniem, kas jau tiek veidoti NVIDIA ekosistēmā. NVIDIA Omniverse, kas līdz 2025. gadam ir kļuvusi par vadošo rīku darbam ar USD formātu - universālu ainas aprakstu, kas ar laiku varētu aizstāt IFC formātu projektēšanas informācijas nodošanas procesos. Kopā ar Isaac Sim, robotizētu procesu simulatoru [163], tādi risinājumi kā NeMo Curator, Cosmos un Omniverse pārstāv jaunu automatizācijas līmeni: no datu tīrīšanas un filtrēšanas līdz mācību kopu ģenerēšanai, objektu īpašību modelēšanai un robotu apmācībai būvlaukumā. Turklat visi šie rīki ir bezmaksas un ar atvērto pirmkodu, kas ievērojami samazina šķēršļus to ieviešanai inženiertehniskajā un būvniecības praksē.

Automātiska datu klasifikācija strukturētu tabulu līmenī nav tik sarežģīts uzdevums, kā tas var šķist pirmajā brīdī. Kā mēs parādījām iepriekšējā nodaļā ( 9.1-10. attēls), ja ir uzkrāti vēsturiski dati, trūkstošās vai nepareizās klases vērtības ir iespējams papildināt, pamatojoties uz līdzīgiem citu elementu parametriem. Ja elementi ar līdzīgiem parametriem jau ir pareizi klasificēti vairākos pabeigtos projektos, sistēma ar lielu varbūtību var ieteikt piemērotu vērtību jaunam vai nepabeigtam elementam ( 10.1-6. att.). Šāda loģika, kas balstīta uz vidējām vērtībām un konteksta analīzi, var būt īpaši efektīva, masveidā apstrādājot tabulāros datus, kas iegūti no tāmēm, specifikācijām vai CAD modeļiem.

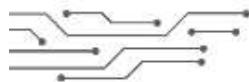


Attēls 10.1-6 Mašīnmācīšanās var palīdzēt automātiski atrast vidējās vērtības neaizpildītiem (baltie lodzini) tabulas parametriem, pamatojoties uz iepriekšējiem projektiem.

Ņemot vērā tik straujo progresu mašīnmācīšanās jomā, ir skaidrs, ka 2025. gadā ir naivi ticēt, ka ēku elementu automātiskas klasificēšanas problēma vēl ilgi nebūs atrisināta. Jā, pašreizējie algoritmi vēl nav pilnībā nobrieduši, jo īpaši attiecībā uz nepilnīgiem vai neviendabīgiem datiem, taču pielāgošanās iespēju logs strauji tuvojas.

Priekšrocībās būs uzņēmumi, kas jau ir ieguldījuši līdzekļus datu vākšanā, attīrīšanā un organizēšanā, kā arī izmanto ETL automatizācijas rīkus. Pārējie riskē atpalikt - gluži tāpat kā uzņēmumi savulaik nespēja tikt galā ar digitālās transformācijas izaicinājumiem transporta un finanšu nozarēs.

Tie, kas turpina paļauties uz manuālu datu pārvaldību un tradicionālām izmaksu un laika aplēšu metodēm, riskē nonākt 2000. gadu taksometru autoparku situācijā, kas līdz 2020. gadu sākumam nespēs pielāgoties mobilo lietotņu un automatizētu maršrutu aprēķinu laikmetam.



## NODAĻA 10.2.

### PRAKTISKS CEĻVEDIS UZ DATIEM BALSTĪTAS PIEEJAS IEVIEŠANAI

#### No teorijas uz praksi: ceļvedis digitālajai transformācijai būvniecībā

Būvniecības nozare pamazām sāk jaunu attīstības posmu, kurā ierastos procesus aizvien vairāk papildina - un dažkārt pat aizstāj - digitālās platformas un pārredzami mijiedarbības modeļi. Tas uzņēmumiem rada ne tikai izaicinājumus, bet arī ievērojamas iespējas. Tās organizācijas, kas jau šodien veido ilgtermiņa digitālo stratēģiju, varēs ne tikai saglabāt savas pozīcijas tirgū, bet arī paplašināt tās, piedāvājot klientiem modernas pieejas un uzticamus, tehnoloģiski pamatotus risinājumus.

Ir svarīgi apzināties, ka zināšanas par koncepcijām un tehnoloģijām ir tikai sākuma punkts. Vadītāji un speciālisti saskaras ar praktisku jautājumu: kur sākt īstenošanu un kā teorētiskās idejas pārvērst reālā vērtībā. Turklāt arvien biežāk rodas jautājums: uz ko balstīties uzņēmējdarbība, ja tradicionālās izmaksu aprēķināšanas un laika grafiku metodes klients jebkurā brīdī var pārskatīt.

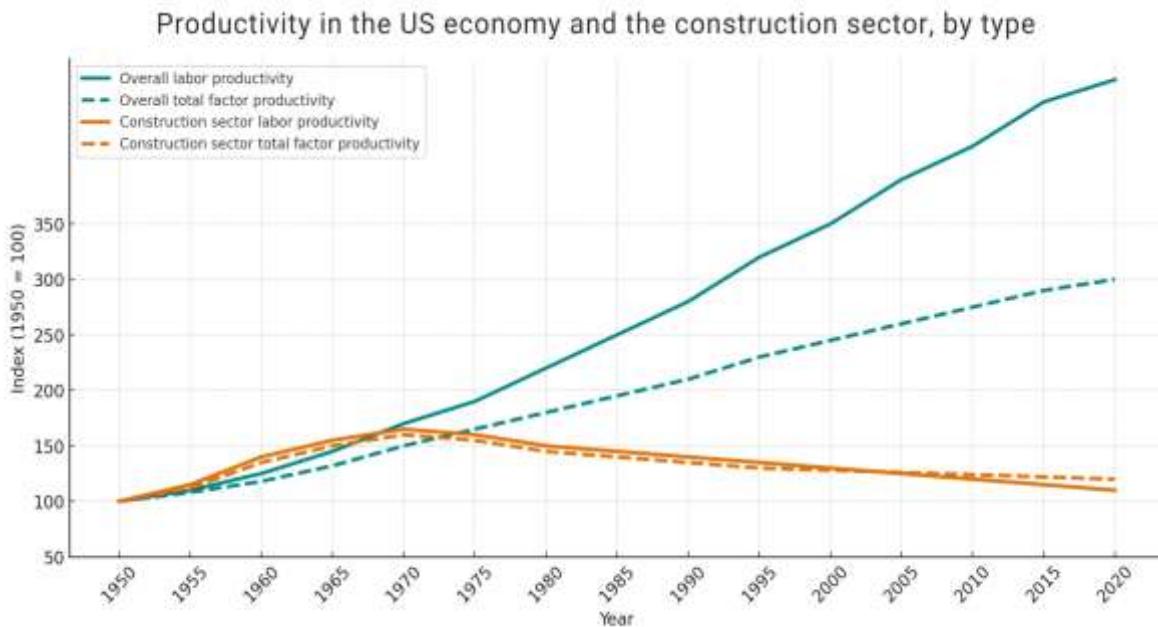
Iespējams, atbilde meklējama ne tik daudz tehnoloģijās, cik jaunas profesionālās kultūras veidošanā, kur darbs ar datiem tiek uztverts kā neatņemama ikdienas prakses sastāvdaļa. Tieši uzmanības trūkums digitālajām tehnoloģijām un inovācijām ir ieaudzinājis būvniecības nozarē nopietnu atpalicību, kas vērojama pēdējās desmitgadēs [43].

Saskaņā ar McKinsey datiem būvniecības nozarē pētniecības un izstrādes izdevumi ir mazāki par 1% no ieņēmumiem, bet automobiļu un aviācijas nozarē tie sasniedz 3,5-4,5%. Līdzīgi arī IT izmaksas būvniecībā joprojām ir mazāk nekā 1% no kopējiem ieņēmumiem [107].

Rezultātā būvniecības nozarē samazinās ne tikai automatizācijas līmenis, bet arī darba ražīgums, un līdz 2020. gadam būvstrādnieks ražos mazāk nekā pirms pusgadsimta ( 10.2-1. attēls)

Šādas produktivitātes problēmas būvniecības nozarē ir raksturīgas lielākajai daļai attīstīto un jaunattīstības valstu (būvniecības produktivitāte ir samazinājusies 16 no 29 ESAO valstīm ( 2.2.2-1. attēls)), un tās norāda ne tikai uz tehnoloģiju trūkumu, bet arī uz nepieciešamību veikt sistēmiskas izmaiņas pašā pieejā vadībai, apmācībai un inovācijām.

Digitālās transformācijas panākumi ir atkarīgi ne tik daudz no rīku skaita un pieejamības, bet gan no organizāciju spējas pārskatīt savus procesus un attīstīt pārmaiņām atvērtu kultūru. Galvenais ir nevis pati tehnoloģija, bet gan cilvēki un procesi, kas nodrošina tās efektīvu izmantošanu, atbalsta nepārtrauktu mācīšanos un veicina jaunu ideju pieņemšanu.



Attēls 10.2-1 Darba ražīguma un kopējās resursu produktivitātes paradokss ASV ekonomikā un būvniecības nozarē (1950-2020) (pamatojoties uz [43]).

Grāmatas pirmajās daļās uzņēmējdarbības vides modelis tika salīdzināts ar meža ekosistēmu (2.1-2. attēls, 1.2-4. attēls, 1.3-2. attēls). Veselā mežā periodiskiem ugunsgrēkiem, neraugoties uz to postošo spēku, ir galvenā loma ilgtermiņa atjaunošanā. Tie attīra augsti no vecās veģetācijas, atgriež uzkrātās barības vielas un rada vietu jaunai dzīvībai. Dažas augu sugas pat ir attīstījušās tā, ka to sēklas atveras tikai tad, kad tās ir pakļautas augstai ugunsgrēka temperatūrai - dabisks mehānisms, kas nodrošina ideālu laiku dīgšanai.

Līdzīgi arī uzņēmējdarbībā krīzes var darboties kā "kontrolēta izdegšana", veicinot jaunu pieeju un uzņēmumu, kas nav piesaistīti novecojušām sistēmām, rašanos. Šādi periodi piespiež atteikties no neefektīvas prakses, atbrīvojot resursus inovācijām. Tāpat kā mežs pēc ugunsgrēka sākas ar pionieru augiem, tā arī uzņēmums pēc krīzes veido jaunus, elastīgus procesus, kas klūst par pamatu nobriedušai informācijas videi.

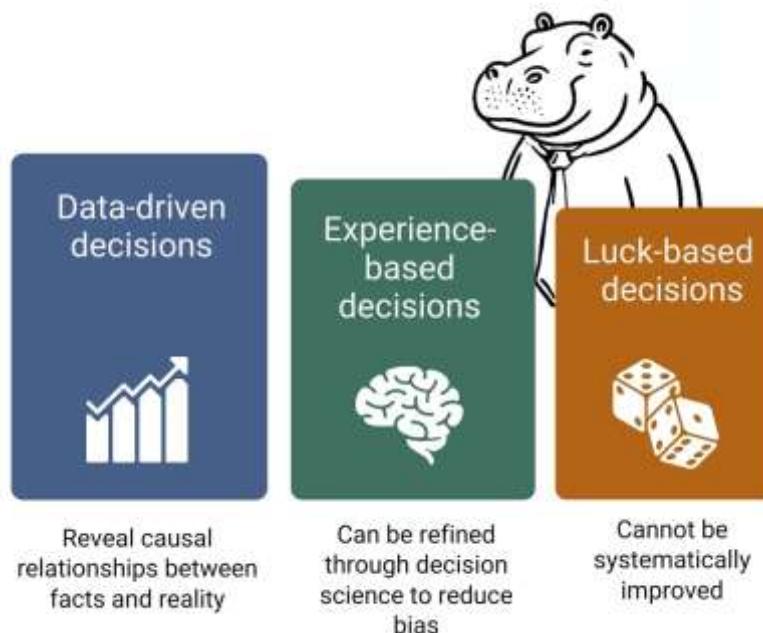
Uzņēmumi, kas spēs pareizi interpretēt šos "signālu ugunsgrēkus" un pārveidot to graujošo energiju konstruktīvās pārmaiņās, sasniegta jaunu veikspējas līmeni - ar pārredzamākiem, adaptīvākiem datu procesiem, kas uzlabo organizācijas dabisko spēju atjaunoties un augt.

Mākslīgā intelekta un mašīnmācīšanās pieaugošā ietekme uz uzņēmējdarbības vidi vairs nav apšaubāma. Tā nav tikai īslaicīga tendence, bet gan stratēģiska nepieciešamība. Uzņēmumi, kas ignorē mākslīgo intelektu, riskē zaudēt konkurētspēju tirgū, kas arvien vairāk veicina inovācijas un elastīgumu.

Nākotne pieder tiem, kas mākslīgo intelektu uzskata ne tikai par rīku, bet arī par iespēju pārdomāt katru uzņēmējdarbības aspektu - no procesu optimizācijas līdz vadības lēmumu pieņemšanai.

## Digitālā pamata veidošana: 1-5 soļi ceļā uz digitālo briedumu

Šajā nodaļā mēs aplūkojam digitālās transformācijas ceļvedi un identificējam galvenos soļus, kas nepieciešami, lai īstenotu uz datiem balstītu pieeju, kas var palīdzēt pārveidot gan uzņēmuma kultūru, gan informācijas ekosistēmu.



Attēls 10.2-2 Kontrolēta atjaunināšana un stratēģijas izvēle: gadījums, pieredze vai dati.

Saskaņā ar McKinsey pētījumu "Why Digital Strategies Fail" (2018) ir vismaz pieci iemesli [164], kāpēc uzņēmumiem neizdodas sasniegt digitālās transformācijas mērķus.

- **Neskaidras definīcijas:** vadītāji un menedžeri atšķirīgi izprot, ko nozīmē "digitālais", un tas rada pārpratumus un pretrunas.
- **Nepietiekama izpratne par digitālo ekonomiku:** daudzi uzņēmumi nepietiekami novērtē digitalizācijas radīto pārmaiņu mērogu uzņēmējdarbības modeļos un nozares dinamikā ( 10.1-6. attēls).
- **Ekosistēmu ignorēšana:** uzņēmumi koncentrējas uz atsevišķiem tehnoloģiju risinājumiem (datu silosiem), ignorējot nepieciešamību integrēties plašākās digitālajās ekosistēmās ( 2.2-2. attēls, 4.1-12. attēls).
- **Konkurentu digitalizācijas nenovērtēšana:** vadītāji nejem vērā, ka arī konkurenti aktīvi ievieš digitālās tehnoloģijas, kas var novest pie konkurences priekšrocību zaudēšanas.
- **Digitalizācijas divējādības iztrūkums:** izpilddirektori deleģē atbildību par digitālo pārveidi citiem vadītajiem, kas birokratizē kontroli un palēnina pārmaiņu procesu.

Lai risinātu šos izaicinājumus, ir nepieciešama skaidra izpratne un digitālo stratēģiju saskaņošana visos organizācijas līmenos. Pirms digitālās stratēģijas izstrādes ir svarīgi saprast, no kā sākt.

Daudzas organizācijas mēdz pieņemt jaunus rīkus un platformas, nesagatavojot pilnīgu priekšstatu par pašreizējo stāvokli.

### **1. solis: Veiciet pašreizējo sistēmu un datu auditu.**

Pirms mainīt procesus, ir svarīgi saprast, kas jau ir ieviests. Audita veikšana ļauj identificēt datu pārvaldības nepilnības un saprast, kādus resursus var izmantot. Audits ir sava veida jūsu uzņēmējdarbības procesu "rentgens". Tas ļauj identificēt riska jomas un noteikt, kuri dati ir kritiski svarīgi jūsu projektam vai uzņēmējdarbībai un kuri ir otršķirīgi.

Galvenās darbības:

- Izveidojiet IT vides karti (Draw.io, Lucidchart, Miro, Visio vai Canva). Uzskaitiet izmantotās sistēmas (ERP, CAD, CAFM, CPM, SCM un citas), kas iesaistītas jūsu procesos un kuras mēs aplūkojām nodaļā "Tehnoloģijas un vadības sistēmas mūsdienu būvniecībā" ( 1.2-4. attēls)
- Novērtējiet katras sistēmas datu kvalitātes jautājumus attiecībā uz datu dublēšanās biežumu, iespējamām trūkstošām vērtībām un formāta neatbilstībām katrā sistēmā.
- Identificējiet "sāpīgos punktus" - vietas, kur procesi var sabojāties vai kur bieži nepieciešama manuāla iejaukšanās - imports, eksports un papildu validācijas procesi.

Ja vēlaties, lai komanda uzticētos pārskatiem, jau no paša sākuma pārliecinieties, ka dati ir pareizi.

Kvalitātes datu audits parādīs, kādi dati:

- Nepieciešama tālāka pilnveidošana (jāveido automātiskie tīrišanas procesi vai papildu transformācijas).
- Tie ir "atkritumi", kas tikai aizsērē sistēmas un no kuriem var atbrīvoties, tos vairs neizmantojot procesos.

Šādu revīziju ir iespējams veikt patstāvīgi. Taču dažkārt ir lietderīgi piesaistīt ārējo konsultantu, jo īpaši no citām nozarēm: svaigs skatījums un neatkarība no būvniecības "īpatnībām" palīdzēs apdomīgi novērtēt esošo situāciju un izvairīties no tipiskajām klūdām, kas saistītas ar aizspriedumiem pret noteiktiem risinājumiem un tehnoloģijām.

### **2. posms: identificēt galvenos datu saskaņošanas standartus.**

Pēc audita ir jāizstrādā kopīgi noteikumi darbam ar datiem. Kā jau tika aplūkots nodaļā "Standarti: no nejaušiem failiem līdz intelīgentam datu modelim", tas palīdzēs novērst datu plūsmas.

Ja nebūs vienota standarta, katra komanda turpinās strādāt "pa savam", un jūs uzturēsiet integrāciju "zoodārzu", kurā dati tiek zaudēti ar katru konvertēšanu.

Galvenās darbības:

■ Izvēlieties datu standartus informācijas apmaiņai starp sistēmām:

- Tabulārajiem datiem var izmantot strukturētus formātus, piemēram, CSV, XLSX vai efektīvākus formātus, piemēram, Parquet.
- Vāji strukturētu datu un dokumentu apmaiņai: JSON vai XML.

■ Apgūt darbu ar datu modeļiem:

- Sāciet ar uzdevumu parametrizēšanu konceptuālā datu modeļa līmenī - kā aprakstīts nodaļā "Datu modelēšana: konceptuālais, loģiskais un fiziskais modelis" ( 4.3-2. attēls).
- Padziļinot iedziļināšanos biznesa procesu loģikā, pārejiet uz prasību formalizēšanu, izmantojot parametrus loģiskajā un fiziskajā modeļī ( 4.3-6. attēls).
- Identificēt galvenās vienības, to atribūtus un attiecības procesos, kā arī vizualizēt šīs attiecības - gan starp vienībām, gan starp parametriem ( 4.3-7. attēls).

■ Izmantojiet regulārās izteiksmes (RegEx), lai validētu un standartizētu datus ( 4.4-7. attēls), kā tika aplūkots nodaļā "Strukturētās prasības un RegEx regulārās izteiksmes". RegEx nav sarežģīts, bet ārkārtīgi svarīgs temats darbā, veidojot prasības fizisko datu modeļu līmenī.

Bez datu līmeņa standartiem un procesu vizualizācijas nav iespējams nodrošināt konsekventu un mērogojamu digitālo vidi. Atcerieties, ka "slikti dati ir dārgi". Un kļūdu izmaksas pieaug, jo projekts vai organizācija kļūst sarežģītāka. Formātu unifikācija, nosaukumu, struktūras un validācijas noteikumu definēšana ir ieguldījums nākotnes risinājumu stabilitātē un mērogojamībā.

### 3. solis. Ieviest DataOps un automatizēt procesus.

Bez skaidri definētās arhitektūras uzņēmumi neizbēgami saskarsies ar atšķirīgiem datiem, kas ietverti atsevišķās informācijas sistēmās. Dati būs neintegrēti, dublēsies vairākās vietās, un to uzturēšana būs dārga.

Iedomājieties, ka dati ir ūdens, bet datu arhitektūra ir sarežģīta cauruļvadu sistēma, kas transportē šo ūdeni no tā uzglabāšanas avota līdz izmantošanas vietai. Tieši datu arhitektūra nosaka, kā informācija tiek apkopota, uzglabāta, pārveidota, analizēta un piegādāta galalietotājiem vai lietojumprogrammām.

DataOps (datu operācijas) ir metodoloģija, kas apvieno datu vākšanu, tīrīšanu, validēšanu un izmantošanu vienotā automatizētā procesu plūsmā, kā mēs detalizēti aprakstījām grāmatas 8. daļā.

Galvenās darbības:

■ Izveidojet un pielāgojet ETL -konveijeri, lai automatizētu procesus:

- Izraksts: organizējiet automātisku datu vākšanu no PDF dokumentiem ( 4.1-2. attēls, 4.1-5. attēls, 4.1-7. attēls), Excel izklājlāpām, CAD modeļiem ( 7.2-4. attēls), ERP sistēmām un citiem avotiem, ar kuriem jūs strādājat.

- Pārveidot: iestatiet automātiskus procesus, lai pārveidotu datus vienotā strukturētā formātā un automatizētu aprēķinus, kas notiks ārpus slēgtām lietojumprogrammām ( 7.2-8. attēls).
- Ielādēt: mēģiniet izveidot automātisku datu augšupielādi uz kopsavilkuma tabulām, dokumentiem vai centralizētām krātuvēm ( 7.2-9. attēls, 7.2-13. attēls, 7.2-16. attēls).
- Automatizēt aprēķinu un QTO (Quantity Take-Off) procesus, kā tas aprakstīts nodaļā "QTO Quantity Take-Off: projekta datu grupēšana pēc atribūtiem":
  - Iestatiet automātisku tilpumu iegūšanu no CAD modeļiem, izmantojot API, spraudņus vai reversās inženierijas rīkus ( 5.2-5. attēls).
  - Izveidojiet noteikumus dažādu klašu elementu grupēšanai pēc atribūtiem tabulu veidā ( 5.2-12. attēls).
  - Mēģiniet automatizēt bieži atkārtotus apjoma un izmaksu aprēķinus ārpus modulārām slēgtām sistēmām ( 5.2-15. attēls).
- Sāciet izmantot Python un Pandas, lai apstrādātu datus, kā mēs to aprakstījām nodaļā "Python Pandas: neaizstājams rīks darbam ar datiem":
  - Lietojiet DataFrame, lai strādātu ar XLSX failiem un automatizētu tabulāro datu apstrādi ( 3.4-6. attēls).
  - Automatizēt informācijas apkopošanu un pārveidošanu, izmantojot dažādas Python bibliotēkas.
  - Izmantojiet LLM, lai vienkāršotu gatavu koda bloku un veselu cauruļvadu rakstīšanu (attēls 7.2-18).
  - Mēģiniet izveidot cauruļvadu Python, kas atrod kļūdas vai saskata anomālijas un nosūta paziņojumu atbildīgajai personai (piemēram, projekta vadītājam) ( 7.4-2. attēls).

Automatizācija, kas balstīta uz DataOps principiem, ļauj pāriet no manuālas un fragmentētas datu apstrādes uz ilgtspējīgiem un atkārtojamiem procesiem. Tas ne tikai samazina slogu darbiniekiem, kuri ikdienā veic vienas un tās pašas transformācijas, bet arī ievērojami palielina visas informācijas sistēmas uzticamību, mērogojamību un pārredzamību.

#### 4. solis: Izveidojiet atvērtu datu pārvaldības ekosistēmu.

Neraugoties uz slēgtu moduļu sistēmu attīstību un to integrāciju ar jauniem rīkiem, uzņēmumi saskaras ar nopietnu problēmu - šādu sistēmu sarežģītības pieaugums pārsniedz to lietderību. Sākotnējā ideja izveidot vienotu patentētu platformu, kas aptvertu visus uzņēmējdarbības procesus, ir novedusi pie pārmērīgas centralizācijas, kur jebkuras izmaiņas prasa ievērojamus resursus un laiku pielāgošanai.

Kā mēs jau aprakstījām nodaļā "Korporatīvais micēlijs: kā dati savieno uzņēmējdarbības procesus", efektīvai datu pārvaldībai ir nepieciešama atvērta un vienota ekosistēma, kas savieno visus informācijas avotus.

Ekosistēmas galvenie elementi:

- Izvēlieties piemērotu datu krātuvi:
  - Tabulām un aprēķiniem izmantojiet datu bāzes, piemēram, PostgreSQL vai MySQL ( 3.1-7. attēls).
  - Dokumentiem un pārskatiem var būt piemērota mākoņglabāšana (Google Drive, OneDrive) vai sistēmas, kas atbalsta JSON formātu.
  - Pārbaudiet datu noliktavu, datu ezeru un citu rīku iespējas centralizētai liela apjoma informācijas glabāšanai un analīzei ( 8.1-8. attēls).
- ieviest risinājumus, lai piekļūtu patentētiem datiem:
  - Ja izmantojat patentētas sistēmas, konfigurējiet piekļuvi tām, izmantojot API vai SDK, lai iegūtu datus ārējai apstrādei ( 4.1-2. attēls).
  - Iepazīstieties ar apgrieztās inženierijas rīku iespējām CAD formātiem ( 4.1-13. attēls).
  - Izveidojiet ETL cauruļvadus, kas periodiski vāc datus no lietojumprogrammām vai serveriem, konvertē tos atvērtos strukturētos formātos un saglabā krātuvēs (attēls 7.2-3).
  - Apspriest komandā, kā nodrošināt piekļuvi datiem, neizmantojot patentētu programmatūru.
  - Atcerieties: dati ir svarīgāki par saskarnēm. Ilgtermiņa vērtību nodrošina informācijas struktūra un pieejamība, nevis konkrēti lietotāja saskarnes rīki.
- Padomājiet par izcilības centra (CoE) izveidi datu jomā, kā tas ir aprakstīts nodaļā "Izcilības centrs (CoE) datu modelēšanai", vai arī par to, kā jūs varat nodrošināt datu ekspertīzi citos veidos (4.3-9. attēls).

Datu pārvaldības ekosistēma rada vienotu informācijas telpu, kurā visi projekta dalībnieki strādā ar konsekventu, aktuālu un pārbaudītu informāciju. Tā ir pamats mērogajamiem, elastīgiem un uzticamiem digitālajiem procesiem

### **Datu potenciāla atraisīšana: 5-10 soļi uz digitālo briedumu**

Papildus tehniskajai integrācijai svarīgs faktors veiksmīgai digitālo risinājumu ieviešanai ir to pieņemšana galalietotāju vidū. Klientu vai lietotāju iesaistišana veikspējas mērišanā ir gan lietotāju pieredzes uzlabošanas, gan pārmaiņu vadības izaicinājums uzņēmumā. Ja risinājums neiekļaujas ierastajā darba plūsmā vai nerisina reālas lietotāju vai klientu problēmas, tas netiks izmantots, un nekādi papildu pasākumi un stimuli to nenovērsīs.

Pārveidošana ir iteratīvs process, kas balstās uz lietotāju mijiedarbības datu analīzi ar jaunajiem procesiem, ar biežiem testēšanas cikliem, pastāvīgu atgriezenisko saiti un pilnveidošanu.

#### **5. solis: Izveidojiet datu kultūru, apmāciet darbiniekus un apkopojiet atgriezenisko saiti.**

Pat vismodernākā sistēma nedarbosies bez darbinieku iesaistīšanās. Jums ir jārada vide, kurā dati tiek izmantoti ikdienā, un komanda saprot to vērtību.

Apvienotās Karalistes valdības publicētajā 2024. gada ziņojumā "Datu analītika un mākslīgais intelekts valdības projektu īstenošanā" norādīts [83], ka, lai sekmīgi īstenotu datu analītiku un mākslīgo intelektu, ir ļoti svarīgi apmācīt speciālistus, kuriem ir nepieciešamās kompetences datu apstrādē un interpretācijā.

Datu analīzes zināšanu trūkums ir viena no galvenajām problēmām, kas ierobežo digitālo transformāciju. Vadītāji ir pieraduši pie iedibinātās rutīnas: ceturkšņa cikliem, prioritārām iniciatīvām un tradicionāliem projektu virzības veidiem. Pārmaiņām ir nepieciešams izteikts līderis - pietiekami augsta ranga, lai viņam būtu ietekme, bet ne tik augsta, lai viņam vai viņai nepietiku laika un motivācijas vadīt ilgtermiņa transformācijas projektu.

Galvenās darbības:

■ Atzīstot nepieciešamību pāriet no subjektīviem lēmumiem, kas balstīti uz augsti atalgota darbinieka viedokli (HiPPO), uz lēmumu pieņemšanas kultūru, kas balstīta uz faktiem un datiem, kā aprakstīts nodaļā "HiPPO jeb viedokļa bīstamība lēmumu pieņemšanā" (attēls 2.1-9).

■ Organizēt sistemātisku apmācību:

- rīkot apmācības par strukturētu datu izmantošanu un pieaicināt ekspertus no citām nozarēm, kuri nav tendēti uz šobrīd būvniecības nozarē populāriem produkiem un koncepcijām.
- Apspriest datu analīzes pieejas un rīkus ar kolēģiem un patstāvīgi apgūt praktisku darbu ar tādiem rīkiem kā Python, pandas un LLM ( 4.1-3. att., 4.1-6. att.).
- Izveidojiet mācību materiālu bibliotēku (vēlamas arī siem videoklipiem) par datu strukturēšanu ( 3.2-15. attēls) un datu modeļu izveidi ( 4.3-6. attēls, 4.3-7. attēls).

■ izmantot modernās mācību tehnoloģijas:

- Valodas modeļu (LLM) izmantošana, lai atbalstītu koda un datu manipulācijas, tostarp koda ģenerēšanu, pārveidošanu un analīzi, kā arī tabulārās informācijas apstrādi un interpretāciju ( 3.4-1. attēls).
- Izpētīt, kā LLM ģenerēto kodu var pielāgot un integrēt pilnīgā cauruļvada risinājumā, strādājot bezsaistes izstrādes vidē (IDE) ( 4.4-14. att., 5.2-13. att.).

Ja vadītājs turpina pieņemt lēmumus "vecmodīgi", nekādas apmācības nepārliecinās cilvēkus uztvert analītiku nopietni.

Datu kultūras veidošana nav iespējama bez nepārtrauktas atgriezeniskās saites. Atgriezeniskā saite ļauj identificēt nepilnības procesos, rīkos un stratēģijās, ko nevar atklāt, izmantojot iekšējos pārskatus vai formālus KPI metriku. Jūsu risinājumu lietotāju komentāri nesniegs praktisku vērtību. Vērtīga ir kritiska atgriezeniskā saite, jo īpaši, ja tā ir balstīta uz konkrētiem novērojumiem un faktiem. Taču šādas informācijas iegūšana prasa pūles: jums ir jāizveido procesi, kuros dalībnieki - gan iekšējie, gan

ārējie - var dalīties ar komentāriem (var būt lietderīgi to darīt anonīmi), neizkroplojot un nebaudoties, ka viņu viedoklis var ietekmēt viņu pašu darbu. Ir svarīgi, lai viņi to darītu bez izkroplojumiem un bez bailēm no negatīvām sekām sev pašiem.

*Jebkura mācīšanās galu galā ir pašmācīšanās [165].*

- Miltons Frīdmens, amerikānu ekonomists un statistiķis

Analītisko rīku ieviešanu jāpapildina ar regulāru to efektivitātes pārbaudi praksē (ROI, KPI), ko var panākt tikai ar strukturētu atgriezenisko saiti no darbiniekiem, klientiem un partneriem. Tas ļauj uzņēmumiem ne tikai izvairīties no kļūdu atkārtošanas, bet arī ātrāk pielāgoties vides izmaiņām. Atgriezeniskās saites apkopošanas un analīzes mehānisma esamība ir viena no organizācijas brieduma pazīmēm, kas liecina par pāreju no gadījuma rakstura digitālajām iniciatīvām uz ilgtspējīgu nepārtrauktu uzlabojumu modeļi ( 2.2-5. attēls).

## 6. posms: No izmēģinājuma projektiem līdz paplašināšanai

*Izvēlieties pietiekami lielas cīņas, lai tās būtu svarīgas, un pietiekami mazas, lai tajās uzvarētu.*

- Džonatans Kozols

Uzsākt digitālo transformāciju "uzreiz un visur" ir ārkārtīgi riskanti. Efektīvāka pieeja ir sākt ar izmēģinājuma projektiem un pakāpeniski paplašināt veiksmīgo praksi.

Galvenās darbības:

■ Izvēlieties pareizo izmēģinājuma projektu:

- Noteikt konkrētu uzņēmējdarbības mērķi vai procesu ar izmērāmiem rezultātiem (KPI, ROI) ( 7.1-5. attēls).
- Izvēlieties ETL automatizācijas procesu, piemēram, automātisku datu validāciju vai darba slodzes aprēķinu (QTO), izmantojot Python un Pandas ( 5.2-10. attēls)
- Noteikt skaidrus panākumu rādītājus (piemēram, samazināt pārbaudes specifikāciju vai datu validācijas ziņojumu sagatavošanas laiku no nedēļas līdz dienai).

■ Izmantojet iteratīvas pieejas:

- Sāciet ar vienkāršiem datu konvertēšanas procesiem un izveidojet daudzformātu datu plūsmas konvertēšanu procesiem nepieciešamajos formātos ( 4.1-2. attēls, 4.1-

5. attēls).

- Pakāpeniski palieliniet uzdevumu sarežģītību un paplašiniet procesu automatizāciju, veidojot pilnīgu cauruļvadu IDE, kas balstīts uz dokumentētiem koda blokiem ( 4.1-7. attēls, 7.2-18. attēls).
- dokumentēt un ierakstīt (vēlams, izmantojot ūsus video) veiksmīgus risinājumus un daļīties ar tiem ar kolēģiem vai profesionālās kopienās.
- Izstrādājiet veidnes un pievienoto dokumentāciju, lai atkārtotu šādus risinājumus un lai tos varētu efektīvi izmantot jūsu kolēģi (vai profesionālās kopienas locekļi, tostarp sociālo plašsaziņas līdzekļu lietotāji).

Pakāpeniska "roll-up" ļauj saglabāt augstu izmaiņu kvalitāti un neiekrist paralēlo ieviešanu haosā. Stratēģija "no maza uz lielu" samazina riskus un ļauj mācīties no mazām klūdām, neļaujot tām izvērsties par kritiskām problēmām.

Pāreja no projektu pieejas, kurā darbinieki ir iesaistīti tikai daļēji, uz pastāvīgu komandu (piemēram, kompetences centru - CoE) veidošanu pašās nodrošināt ilgtspējīgu produkta attīstību arī pēc tā pirmās versijas izlaišanas. Šādas komandas ne tikai atbalsta esošos risinājumus, bet arī turpina tos uzlabot.

Tas samazina atkarību no ilgstošas apstiprināšanas: komandas locekļi ir pilnvaroti pieņemt lēmumus savas atbildības jomā. Rezultātā vadītāji ir atbrīvoti no nepieciešamības veikt mikrolīmenedžēšanu, un komandas var koncentrēties uz reālas vērtības radīšanu.

Jaunu risinājumu izstrāde nav sprints, bet gan maratons. Tie, kam tas izdodas, ir tie, kas sākotnēji tiecas uz ilgtermiņa un konsekventu darbu.

Ir svarīgi apzināties, ka tehnoloģijām ir nepieciešama pastāvīga attīstība. Ieguldījumi tehnoloģisko risinājumu ilgtermiņa attīstībā ir veiksmīgas darbības pamats.

## 7. solis: izmantojiet atvērtus datu formātus un risinājumus

Kā jau minēts nodaļās par moduļu platformām (ERP, PMIS, CAFM, CDE u. c.), ir svarīgi pievērst uzmanību atvērtiem un universāliem datu formātiem, kas nodrošina neatkarību no piegādātāju risinājumiem un palielina informācijas pieejamību visām ieinteresētajām personām.

Galvenās darbības:

- Pāreja no slēgtiem formātiem uz atvērtiem formātiem:
  - Izmantojiet atvērtus formātus, nevis patentētus formātus, vai atrodiet veidu, kā iestatīt automātisku slēgto formātu augšupielādi vai pārveidošanu atvērtos formātos ( 3.2-15. attēls).
  - Ieviest rīkus darbam ar Parquet, CSV, JSON, XLSX, kas ir apmaiņas standarti starp lielāko daļu moderno sistēmu (8.1-2. attēls.).
  - Ja jūsu procesos svarīga nozīme ir darbam ar 3D ģeometriju, apsveriet iespēju izmantot atvērtus formātus, piemēram, USD, glTF, DAE vai OBJ ( 3.1-14. attēls).

■ Izmantojiet vektoru datubāzes datubāzes, lai efektīvi analizētu un meklētu informāciju:

- Izmantojiet Bounding Box un citas metodes, lai vienkāršotu 3D ģeometriju ( 8.2-1. attēls).
- Pārdomājiet, kur varat īstenot datu vektorizāciju - pārvērst tekstus, objektus vai dokumentus skaitliskos attēlos ( 8.2-2. attēls).

■ Lielo datu analīzes rīku izmantošana:

- Organizēt uzkrāto vēsturisko datu (piemēram, PDF, XLSX, CAD) glabāšanu analīzei piemērotos formātos (Apache Parquet, CSV, ORC) ( 8.1-2. att.).
- Sākt pielietot statistikas pamatmetodes un strādāt ar reprezentatīvām izlasēm vai vismaz iepazīties ar statistikas pamatprincipiem ( 9.2-5. attēls).
- Ieviest un apgūt rīkus datu vizualizēšanai un datu savstarpējām attiecībām, lai vizualizētu analīzes rezultātus. Bez labas vizualizācijas nav iespējams pilnībā izprast ne pašus datus, ne uz tiem balstītos procesus ( 7.1-4. attēls).

Pāreja uz atvērtiem datu formātiem un rīku ieviešana informācijas analīzei, glabāšanai un vizualizēšanai ir pamats ilgtspējīgai un neatkarīgai digitālajai pārvaldībai. Tas ne tikai samazina atkarību no piegādātājiem, bet arī nodrošina vienlīdzīgu piekļuvi datiem visām ieinteresētajām personām.

## 8. posms. Sāciet īstenot prognozēšanas mašīnmācīšanos

Daudzi uzņēmumi ir uzkrājuši milzīgus datu apjomus - sava veida "informācijas geizerus", kas joprojām netiek izmantoti. Šie dati ir savākti simtos vai tūkstošos projektu, bet bieži vien turpmākajos procesos ir izmantoti tikai vienu reizi vai nav izmantoti vispār. Slēgtos formātos un sistēmās glabātie dokumenti un modeļi bieži tiek uztverti kā novecojis un nederīgs balasts. Tomēr patiesībā tie ir vērtīgs resurss - pamats, uz kura analizēt pieļautās klūdas, automatizēt rutīnas darbības un izstrādāt inovatīvus risinājumus automātiskai klasifikācijai un pazīmju atpazīšanai turpmākajos projektos.

Galvenais izaicinājums ir iemācīties iegūt šos datus un pārveidot tos praktiski izmantojamās atzinās. Kā aprakstīts nodaļā par mašīnmācīšanos un prognozēm, mašīnmācīšanās metodēm ir potenciāls ievērojami uzlabot aplēšu un prognožu precizitāti dažādos ar būvniecību saistītos procesos. Pilnībā izmantojot uzkrātos datus, paveras celš uz efektivitātes uzlabošanu, riska samazināšanu un ilgtspējīgu digitālo procesu izveidi.

Galvenās darbības:

■ Sāciet ar vienkāršiem algoritmiem:

- Mēģiniet piemērot lineāro regresiju, izmantojot LLM ieteikumus, lai prognozētu atkārtotu veikspēju datu kopās, kurās nav vai ir minimāla atkarība no daudziem faktoriem ( 9.3-4. attēls).
- Apsveriet, kuros jūsu procesu posmos teorētiski varētu izmantot k-tuvāko kaimiņu (k-NN) algoritmu - piemēram, klasifikācijas uzdevumiem, objektu līdzības novērtēšanai vai prognozēšanai, pamatojoties uz vēsturiskiem analogiem ( 9.3-5. attēls).

■ Vākt un strukturēt datus, lai apmācītu modeļus:

- Apkopot vēsturiskos projekta datus vienuviet un vienotā formātā (attēls 9.1-10).
- Darbs pie mācību paraugu kvalitātes un reprezentativitātes, izmantojot automatizētu ETL (9.2.2-8. attēls).
- Iemācīties sadaļīt datus mācību un testa kopās, kā mēs to darījām "Titānika" datu kopas piemērā ( 9.2-6. attēls, 9.2-7. attēls).

■ Apsveriet iespēju paplašināt mašīnmācīšanās metožu pielietojumu visdažādākajiem uzdevumiem, sākot no projektu grafiku prognozēšanas līdz loģistikas optimizācijai, resursu pārvaldībai un potenciālo problēmu agrīnai identificēšanai.

Mašīnmācīšanās ir rīks, ar kura palīdzību arhivētus datus var pārvērst par vērtīgu aktīvu prognozēšanai, optimizācijai un uz informāciju balstītu lēmumu pieņemšanai. Sāciet ar nelielām datu kopām ( 9.2-5. attēls) un vienkāršiem modeļiem, pakāpeniski palielinot to sarežģītību.

## 9. posms. Integrējet IoT un modernās datu vākšanas tehnoloģijas.

Būvniecības pasaule strauji klūst digitāla: katrs būvniecības fotoattēls, katrs Teams ieraksts jau ir daļa no plašāka realitātes parametrizēšanas un marķēšanas procesa. Tāpat kā GPS savulaik pārveidoja loģistiku, IoT, RFID un automātiskās datu vākšanas tehnoloģijas maina būvniecības nozari. Kā aplūkots nodaļā "Lietu internets un viedie līgumi", digitālā būvlaukuma ar sensoriem un automātisko uzraudzību ir nozares nākotne.

Galvenās darbības:

■ Izveist IoT -ierīces, RFID -tags un detalizēti aprakstīt ar tiem saistītos procesus:

- Izvērtējiet, kurās projekta jomās vai posmos var gūt vislielāko ieguldījumu atdevi (ROI), piemēram, temperatūras, vibrācijas, mitruma vai kustības uzraudzībai.
- Apsveriet iespēju izmantot RFID, lai izsekotu materiālus, darbarīkus un aprīkojumu visā piegādes kēdē.
- Apsveriet, kā savāktos datus var integrēt vienotā informācijas sistēmā, piemēram, Apache NiFi, automatizētai apstrādei un analīzei reālajā laikā ( 7.4-5. attēls).

■ Izveidot reālā laika uzraudzības sistēmu:

- Izstrādāt paneļus, lai sekotu galvenajiem procesu vai projektu rādītājiem, izmantojot vizualizācijas rīkus, piemēram, Streamlit, Flask vai Power BI).
- Iestatiet automātiskus paziņojumus, lai signalizētu par kritiskām novirzēm no plāna vai normām ( 7.4-2. attēls).
- Novērtējiet iekārtu prognozējamās tehniskās apkopes potenciālu, pamatojoties uz apkopotajiem datiem un noteiktajiem modeļiem ( 9.3-6. attēls).

■ Apvienojet datus no dažādiem avotiem:

- Sāciet ar datu modeļa vizualizāciju fiziskajā līmenī - atspoguļojiet informācijas

plūsmu struktūru un galvenos parametrus, kas nāk no CAD sistēmām, IoT ierīcēm un ERP platformām ( 4.3-1. attēls).

- Sāciet ar vienotas datu analīzes un vadības lēmumu atbalsta platformas aptuvenu plānu. Apkopojet galvenās funkcijas, datu avotus, lietotājus un paredzamos lietojumu scenārijus ( 4.3-7. attēls).

Jo ātrāk sāksiet savienot reālos procesus ar digitālo pasauli, jo ātrāk varēsiet tos pārvaldīt ar datiem - efektīvi, pārredzami un reālajā laikā.

## 10. solis. Sagatavoieties nākotnes pārmaiņām nozarē

Būvniecības uzņēmumi pastāvīgi saskaras ar ārējās vides spiedienu: ekonomikas krīzes, tehnoloģiju lēcieni, regulējuma izmaiņas. Līdzīgi kā mežs, kuram ir jāiztur lietus, sniegs, sausums un dedzinoša saule, uzņēmumi dzīvo nepārtrauktas pielāgošanās apstākļos. Un tāpat kā koki kļūst noturīgi pret salu un sausumu, pateicoties dzīlai sakņu sistēmai, tikai tās organizācijas, kurām ir stabils automatizētu procesu pamats, spēja paredzēt pārmaiņas un elastīgi pielāgot stratēģijas, saglabā dzīvotspēju un konkurētspēju.

Kā minēts nodaļā "Izdzīvošanas stratēģijas: konkurētspējas priekšrocību veidošana", būvniecības nozarē sākas radikālas pārveides posms. Pasūtītāja un būvuzņēmēja mijiedarbība virzās uz pārliecināšanas modeli, kurā pārredzamība, paredzamība un digitālie rīki aizstāj tradicionālās pieejas. Šajā jaunajā realitātē uzvar nevis lielākie, bet gan elastīgākie un tehnoloģiski attīstītākie.

Galvenās darbības:

- Analizēt uzņēmējdarbības ievainojamības atvērto datu kontekstā:
  - Novērtējiet, kā piekļuves datiem demokratizācija kā daļa no Uberizācijas varētu postoši ietekmēt jūsu konkurences priekšrocības un uzņēmējdarbību ( 10.1-5. attēls).
  - Padomājiet par stratēģiju, kā pāriet no nepārskatāmiem un nošķirtiem procesiem uz uzņēmējdarbības modeļiem, kas balstīti uz atvērtiem risinājumiem, sistēmu sadarbspēju un datu pārredzamību ( 2.2-5. attēls).
- Izstrādājiet ilgtermiņa digitālo stratēģiju:
  - Nosakiet, vai vēlaties būt inovāciju līderis, vai arī dodat priekšroku scenārijam, kurā taupīsiet savus resursus.
  - Aprakstiet posmus: īstermiņa (procesu automatizācija, datu strukturēšana), vidēja termiņa (LLM un ETL ieviešana), ilgtermiņa (digitālās ekosistēmas, centralizēti repozitoriji).
- Padomājiet par pakalpojumu portfeļa paplašināšanu:
  - Apsveriet iespēju piedāvāt jaunus pakalpojumus (koncentrējoties uz energoefektivitāti, ESG, datu pakalpojumiem). Par jauniem uzņēmējdarbības modeļiem plašāk runāsim nākamajā nodaļā.
  - Mēģiniet sevi pozicionēt kā uzticamu tehnoloģiju partneri, kas atbalsta visu objekta

dzīves ciklu - no projektēšanas līdz ekspluatācijai. Uzticība jums jābalsta uz sistemātisku pieeju, pārredzamiem procesiem un spēju nodrošināt ilgtspējīgus tehnoloģiskos risinājumus.

Transformācijas vidē uzvar nevis tie, kas vienkārši reaģē uz pārmaiņām, bet gan tie, kas ir proaktīvi. Elastīgums, atvērtība un digitālā brieduma pakāpe ir ilgtspējības pamats rītdienas veidošanā.

## **Transformācijas ceļvedis: no haosa līdz uz datiem balstītam uzņēmumam**

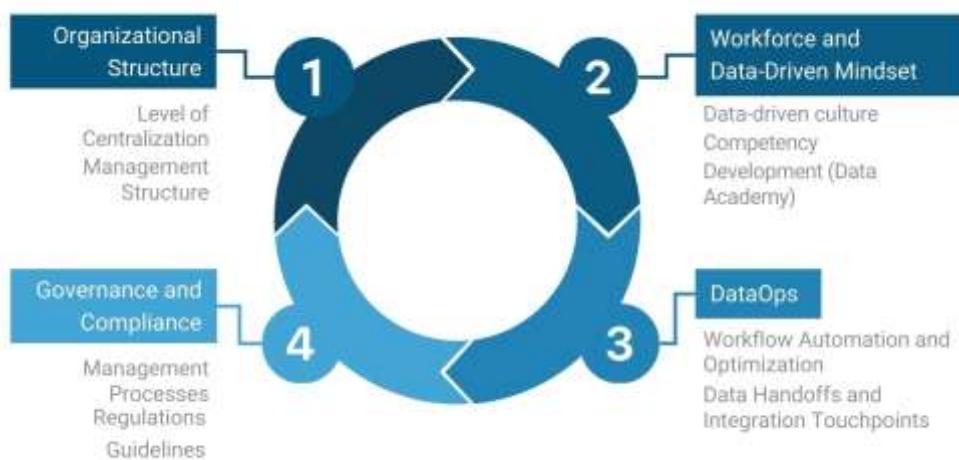
Tālāk sniegtais plāns var kalpot kā sākotnējais kritērijs - sākumpunkts, lai veidotu savu uz datiem balstītu digitālās transformācijas stratēģiju:

- **Audits un standarti:** pašreizējā stāvokļa analīze, datu unifikācija
- **Datu strukturēšana un klasifikācija:** automatizēt nestrukturētu un nesakārtotu datu pārveidošanu.
- **Automatizēt grupēšanu, aprēķinus un aprēķinus:** automatizācijai izmantojiet atvērtā koda rīkus un bibliotēkas.
- **Ekosistēma un COE:** izveidot iekšējo komandu, lai uzņēmumā izveidotu vienotu datu ekosistēmu.
- **Kultūra un mācīšanās:** pāreja no HiPPO -risinājumiem uz uz datiem balstītiem risinājumiem.
- **Izmēģinājuma projekti, atgriezeniskā saite un mēroga palielināšana:** rīkojieties iteratīvi: izmēģiniet jaunas metodes ierobežotā mērogā, iegūstiet derīgu atgriezenisko saiti un pakāpeniski palieliniet risinājumus.
- **Atvērtie formāti:** izmantojiet universālus un atvērtus formātus, lai nodrošinātu neatkarību no programmatūras piegādātājiem.
- **Mašīnmācīšanās:** iestrādāt ML algoritmus prognozēšanas un optimizācijas procesos.
- **IoT un digitālā būvlaukumā:** integrēt modernās datu vākšanas tehnoloģijas procesos
- **Stratēgiskā pielāgošanās:** sagatavoties nākotnes pārmaiņām nozarē

Vissvarīgākais ir atcerēties, ka "dati vien nemaina uzņēmumu: to maina cilvēki, kas zina, kā strādāt ar šiem datiem". Uzsveriet kultūru, pārredzamus procesus un apņemšanos nepārtraukti pilnveidoties.

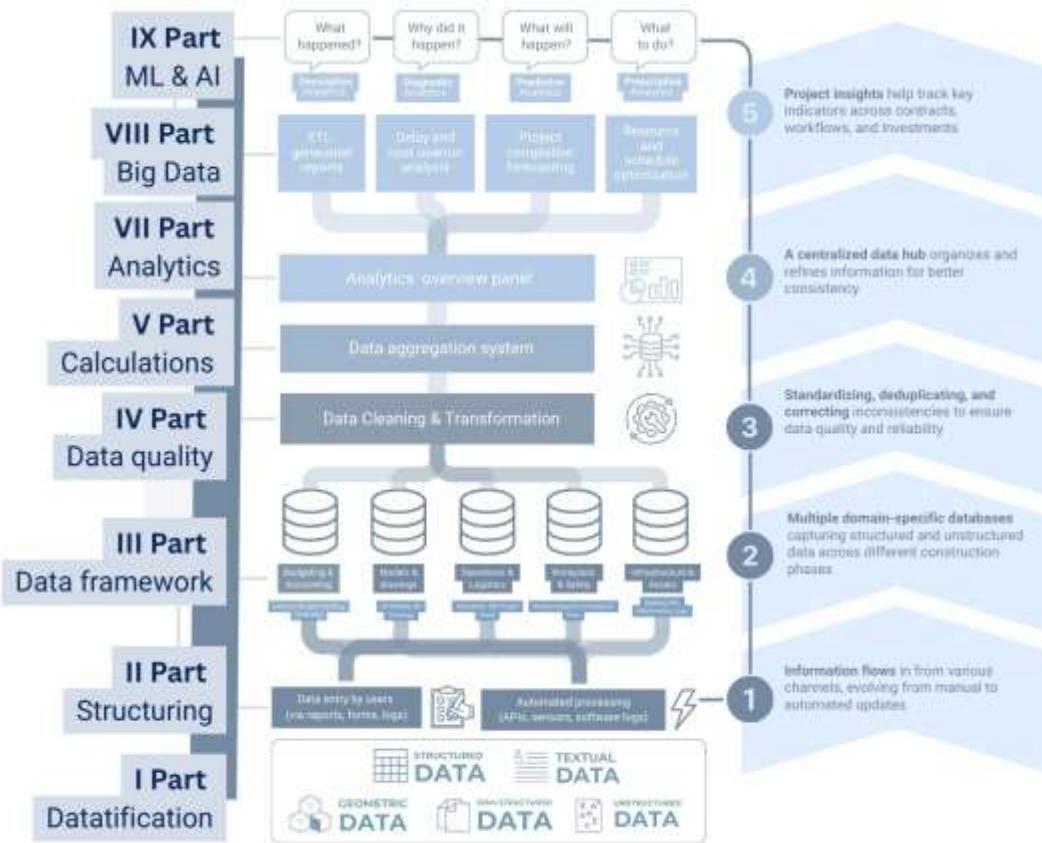
Sistēmu pieeja ļauj pāriet no atsevišķām digitālām iniciatīvām uz pilnvērtīgu, uz datiem balstītu pārvaldības modeli, kurā lēmumi tiek pieņemti, balstoties nevis uz intuīciju vai pieņēmumiem, bet gan uz datiem, faktiem un matemātiski aprēķinātām varbūtībām. Būvniecības nozares digitālā transformācija nav saistīta tikai ar tehnoloģiju ieviešanu, bet gan ar uzņēmējdarbības ekosistēmas veidošanu, kurā projekta informācija tiek nepārtraukti un iteratīvi pārsūtīta starp dažādām sistēmām. To darot, mašīnmācīšanās algoritmi nodrošina automātisku, nepārtrauktu analīzi, prognozēšanu un procesu optimizāciju. Šādā vidē spekulācijas un slēptie dati kļūst nebūtiski - paliek tikai pārbaudīti

modeli, pārskatāmi aprēķini un paredzami rezultāti.



Attēls 10.2-3 Veiksmīgas uzņēmuma līmeņa datu pārvaldības galvenie elementi.

Katra grāmatas daļa atbilst konkrētam datu apstrādes un analīzes posmam būvniecības projektos (2.2-5. attēls). Ja vēlaties atgriezties pie kādas no iepriekš aplūkotajām tēmām un aplūkot to no visaptverošas datu izmantošanas plūsmas izpratnes, varat atsaukties uz daļu nosaukumiem 10.2-4. attēlā.



Attēls 10.2-4 Grāmatas daļas datu apstrādes konveijera kontekstā (attēls 2.2-5): no informācijas digitalizācijas līdz analītikai un mākslīgajam intelektam.

Neatkarīgi no jūsu organizācijas lieluma, tehnoloģiju brieduma līmeņa vai budžeta, jūs varat sākt pāreju uz uz datiem balstītu pieeju jau šodien. Pat mazi soli pareizajā virzienā laika gaitā dos rezultātus.

Uz datiem balstīta transformācija nav vienreizējs projekts, bet gan nepārtraukts, atkārtots uzlabošanas process, kas ietver jaunu rīku ieviešanu, procesu pārveidošanu un uz datiem balstītu lēmumu pieņemšanas kultūras veidošanu.

## Būvniecība 5.0 industrijā: kā pelnīt naudu, kad vairs nav iespējams slēpties

Ilgu laiku būvniecības uzņēmumi ir pelnījuši naudu no procesu nepārredzamības. Galvenais uzņēmējdarbības modelis bija spekulācijas - materiālu izmaksu, darbu apjoma un procentuālo uzcenojumu pārvērtēšana slēgtās ERP - un PMIS - sistēmās, kas nav pieejamas ārējiem auditiem. Klientu un viņu uzticības personu ierobežotā piekļuve sākotnējiem projekta datiem radīja pamatu shēmām, kurās aprēķinu ticamību pārbaudīt kļuva gandrīz neiespējami.

Tomēr šis modelis strauji zaudē aktualitāti. Līdz ar datu pieejamības demokratizāciju, LLM parādīšanos, atvērto datu un ETL automatizācijas rīku parādīšanos nozare pāriet uz jaunu darba standartu.

Rezultātā necaurspīdīgums vairs nav konkurences priekšrocība - drīz tas kļūs par slogu, no kura būs grūti atteikties. Pārredzamība no izvēles iespējas kļūst par priekšnoteikumu, lai noturētos tirgū.

Ar ko klienti - bankas, investori, fiziskie klienti, privātā kapitāla, valsts klienti - sadarbosies jaunajā digitālajā realitātē? Atbilde ir acīmredzama: ar tiem, kas spēj nodrošināt ne tikai rezultātu, bet arī katras soļa ceļā uz to pamatojumu. Pieaugot atvērto datu apjomam, partneri un klienti izvēlēsies uzņēmumus, kas garantē pārredzamību, precīzitāti un rezultātu paredzamību.

Šajā kontekstā rodas jauni uzņēmējdarbības modeļi, kas balstās uz datu pārvaldību un uzticēšanos, nevis spekulācijām:

- **Pārdod procesus, nevis kvadrātmetrus:** galvenais ieguvums ir uzticēšanās un efektivitāte, nevis konkrēti līgumi ar atlaidēm. Galvenā vērtība būs rezultātu paredzamība, pamatojoties uz uzticamiem un pārbaudītiem datiem. Mūsdienīgi uzņēmumi nepārdos būvniecības objektu kā tādu, bet gan:
  - precīzi termiņi un pārskatāmi darba grafiki;
  - pamatotas aplēses, kas pamatotas ar aprēķiniem;
  - pilnīgu digitālu izsekojamību un kontroli visos projekta posmos.
- **Inženierzinātnes un analītika kā pakalpojums:** "Datu kā pakalpojuma" modelis (veids, kā lietotājam internetā piegādāt gatavus datus kā pakalpojumu), kur katrs projekts kļūst par daļu no digitālās datu ķēdes, un uzņēmējdarbības vērtība ir spēja pārvaldīt šo ķēdi. Uzņēmumi pārveidojas par viedām platformām, kas piedāvā risinājumus, kuru pamatā ir automatizācija un analīze:
  - Automatizēta un pārskatāma tāmju un plānu sagatavošana;
  - riska un laika novērtējums, pamatojoties uz mašīnmācīšanās algoritmiem;
  - vides rādītāju (ESG, CO<sub>2</sub>, energoefektivitāte) aprēķināšana;
  - pārskatu sagatavošana no revidētiem atklātiem avotiem.
- **Inženiertehniskās kompetences produktalizācija:** uzņēmuma izstrādājumus var atkārtoti

izmantot uzņēmumā un izplatīt kā atsevišķu produktu, veidojot papildu ienākumu avotu, izmantojot digitālos pakalpojumus. Jaunajā vidē uzņēmumi rada ne tikai projektus, bet arī digitālos aktīvus:

- komponentu bibliotēkas un tāmes veidnes;
- automatizēti verifikācijas moduli;
- Atvērtā koda spraudņi un skripti (konsultāciju pārdošana) darbam ar datiem.

■ **Jauns uzņēmuma veids - uz datiem orientēts integrators:** tirgus dalībnieks, kas nav atkarīgs no konkrētiem programmatūras piegādātājiem vai modulārām sistēmām un nav "ieslēgts" vienā programmatūras saskarnē. Tas brīvi darbojas ar datiem un, balstoties uz tiem, veido savu konkurētspēju. Nākotnes būvniecības uzņēmums ir ne tikai būvuzņēmējs, bet arī informācijas integrators, kas klientam spēj veikt šādas funkcijas:

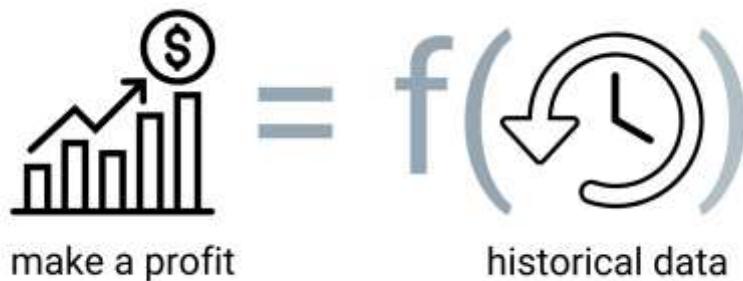
- Apvienojiet datus no dažādiem avotiem un veiciet analīzi;
- Nodrošināt procesu pārredzamību un uzticamību;
- konsultēt par uzņēmējdarbības procesu optimizāciju;
- izstrādāt rīkus, kas darbojas atvērto datu ekosistēmā, LLM, ETL un Pipelines.

Rūpniecība 5.0 ( 2.1-12. attēls) iezīmē "manuālo vidējo koeficientu" un vakara tikšanos starp izpilddirektoriem un aprēķinu un grāmatvedības nodalju beigas. Viss, kas iepriekš bija apslēpts - aprēķini, aplēses, apjomī - kļūst atklāts, pārbaudāms un saprotams pat tiem, kas nav eksperti. Uzvarētāji būs tie, kas pirmie pārorientēsies. Visi pārējie paliks ārpus būvniecības nozares jaunās digitālās ekonomikas.

## SECINĀJUMS

Būvniecības nozarē sākas būtisku pārmaiņu ēra. Sākot ar pirmajiem ierakstiem uz māla plāksnītēm un beidzot ar milzīgajiem digitālo datu apjomiem, kas plūst no projektu serveriem un būvlaukumiem, būvniecības informācijas pārvaldības vēsture vienmēr ir atspoguļojusi sava laika tehnoloģiju briedumu. Šodien, līdz ar automatizācijas, atvērto formātu un inteliģento analīzes sistēmu parādīšanos, nozarē notiek nevis pakāpeniska evolūcija, bet gan strauja digitālā transformācija.

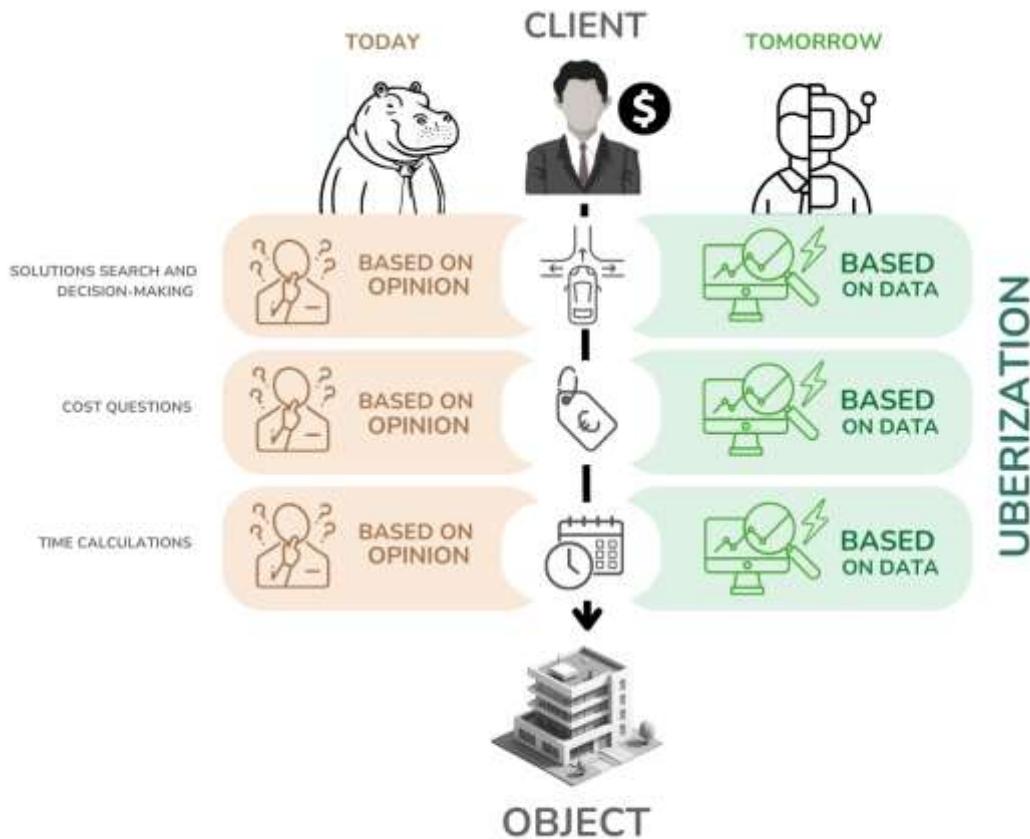
Tāpat kā citās tautsaimniecības nozarēs, arī būvniecībā būs jāpārdomā ne tikai instrumenti, bet arī darba principi. Uzņēmumi, kas agrāk diktēja tirgu un bija galvenais starpnieks starp klientu un projektu, zaudē savu unikālo pozīciju. Uzticēšanās un spēja strādāt ar datiem - no to vākšanas un strukturēšanas līdz lēmumu analīzei, prognozēšanai un automatizācijai - izvirzās priekšplānā.



Attēls 10.2-1. Strukturēti vēsturiskie dati ir degviela efektīvai un vadāmai uzņēmējdarbībai.

Šajā grāmatā detalizēti izklāstīti galvenie datu pārvaldības principi būvniecības nozarē - no audita un standartizācijas līdz procesu automatizācijai, vizualizācijas rīku izmantošanai un inteliģento algoritmu ieviešanai. Mēs aplūkojām, kā pat ar ierobežotiem resursiem var izveidot strādājošu datu arhitektūru un sākt pieņemt lēmumus, pamatojoties uz pārbaudāmiem faktiem, nevis intuīciju. Darbs ar datiem vairs nav tikai IT nodalas uzdevums - tas kļūst par vadības kultūras pamatu, kas nosaka uzņēmuma elastību, pielāgošanās spēju un ilgtermiņa ilgtspēju.

Mašīnmācīšanās tehnoloģiju, automātiskās apstrādes sistēmu, digitālo dvīņu un atvērto formātu izmantošana jau šodien ļauj novērst cilvēcisko faktoru tur, kur tas iepriekš bija ļoti svarīgs. Būvniecība virzās uz autonomiju un kontrolējamību, kur virzību no idejas līdz projekta realizācijai var salīdzināt ar navigāciju autopilota režīmā: bez atkarības no subjektīviem lēmumiem, bez nepieciešamības katrā posmā iejaukties manuāli, bet ar pilnīgu digitālu izsekojamību un kontroli (10.2-2. att.).



Attēlā: 10.2-2 Pāreju no lēmumu pieņemšanas, kas balstīta uz svarīgu ekspertu viedokļiem (HiPPO), uz datu analīzi galvenokārt veicinās klients.

Apgūstot šajā grāmatā aprakstītās metodes, principus un rīkus, jūs varēsiet sākt pieņemt uz datiem, nevis intuīciju balstītus lēmumus savā uzņēmumā. Jūs arī spēsiet palaist moduļu kēdes LLM, kopēt izstrādes vidē (IDE) jau gatavus ETL cauruļvadus un automātiski apstrādāt datus, lai iegūtu vajadzīgo informāciju vēlamajā formā. Balstoties uz grāmatas nodajām par lielajiem datiem un mašīnmācīšanos, jūs pēc tam varēsiet īstenot sarežģītākus scenārijus - iegūt jaunas atziņas no vēsturiskiem datiem un piemērot mašīnmācīšanās algoritmus, lai prognozētu un optimizētu procesus.

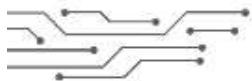
Atklāti dati un procesi nodrošinās pamatu precīzākām projektu izmaksu un termiņu aplēsēm, neļaujot būvniecības uzņēmumiem spekulēt ar neskaidriem datiem. Tas ir gan izaicinājums, gan iespēja nozarei pārdomāt savu lomu un pielāgoties jaunajai videi, kurā pārredzamība un efektivitāte kļūs par galvenajiem veiksmes faktoriem.

Lai gūtu panākumus digitālās transformācijas laikmetā, galvenais ir gatavība pārņemt zināšanas un īstenot tās praksē.

Uzņēmumi, kas to sapratīs pirmie, gūs priekšrocības jaunajā digitālajā konkurencē. Taču ir svarīgi saprast, ka dati vien neko nemainīs. Daudziem cilvēkiem būs jāmaina domāšanas veids, un tam ir nepieciešams stimuls. Jūsu uzņēmumam ir jāpārdomā, kā tas dalās ar datiem

Cilvēki, kas maina uzņēmumu, ir tie, kuri zina, kā strādāt ar šiem datiem, interpretēt tos, izmantot, lai optimizētu un izveidotu jaunu procesu arhitektūru, pamatojoties uz tiem.

Ja lasāt šīs rindas, jūs esat gatavi pārmaiņām un jau esat soli priekšā. Paldies, ka izvēlējāties šo ceļu. Laipni lūgti digitālās transformācijas laikmetā!



## PAR AUTORU

Mani sauc Artjoms Boiko. Mans ceļš būvlaukumā sākās 2007. gadā - strādājot par kalnraču degslānekļa raktuvē savā dzimtajā pilsētā, kamēr studēju Sanktpēterburgas Kalnrūpniecības universitātē, specializējoties raktuvju un pazemes būvniecībā. Uz šīs grāmatas aizmugurējā vāka redzams sprādzienizmetējs sejā, kurā mēs raktuvēs raktuvēs raktuvēs un uzspridzinājām simtiem degslānekļa kubikus. Mana karjera mani ir aizvedusi dažādos virzienos, sākot no kalnraktuvju un pazemes būvdarbu strādnieka līdz rūpnieciskajam alpīnistam, jumta klājējam un lifta iekārtu uzstādītājam. Man ir bijis tas gods piedalīties dažāda mēroga projektos, sākot ar privātmāju būvniecību un beidzot ar lielu rūpniecisku objektu būvniecību dažādos pasaules reģionos.



Laika gaitā mans darbs ir pārorientējies no fiziskās būvniecības uz informācijas pārvaldību un digitālajiem procesiem. Kopš 2013. gada esmu strādājis dažādos amatos mazos, vidējos un lielos būvniecības uzņēmumos vairākos Vācijas reģionos, sākot no projektētāja līdz datu pārvaldības vadītājam. Runājot par datu pārvaldību, mana pieredze ietver darbu ar datiem dažādās ERP sistēmās, CAD (BIM), MEP, FEM, CMS. Esmu bijis iesaistīts optimizācijā, procesu automatizācijā, kā arī analīzē, mašīnmācībā, datu apstrādē būvdarbu plānošanas, aprēķinu un izpildes posmos rūpniecības, dzīvojamo ēku, infrastruktūras un komunālo pakalpojumu būvniecības uzņēmumos.

Ar atvērtā koda programmatūru un atvērtajiem datiem strādāju kopš 2003. gada. Šajā laikā esmu īstenojis daudzus tīmekļa projektus - no tīmekļa vietnēm un tiešsaistes veikaliem līdz pilnvērtīgām tīmekļa lietojumprogrammām, izmantojot atvērtā koda risinājumus un atvērtās CMS. Šīm platformām, kas daudzējādā ziņā ir līdzīgas mūsdienīgām ēku ERP, ir modulāra arhitektūra, augsta pielāgošanās spēja un pieejamība. Šī pieredze lika pamatus manai profesionālajai pieejai - uzsvars uz atvērtā pirmkoda tehnoloģijām un kopīgas izstrādes kultūru. Cieņu pret atvērtā koda programmatūru un brīvu zināšanu apmaiņu es cenšos veicināt būvniecības nozarē. Mans darbs, lai uzlabotu datu pieejamību būvniecības nozarē, ir rezultējies vairāku sociālo mediju kopienu izveidē, lai apspriestu datu atvērtību un atvērtā koda izmantošanu būvniecībā, kā arī vairāku jaunuzņēmumu izveidē, kas izstrādā risinājumus, lai nodrošinātu piekļuvi datiem no dažādām slēgtām sistēmām un platformām.

Mans ieguldījums profesionālu kopienā ir dalība kā referentam konferencēs, kas veltītas CAD sadarbīspējai (BIM), ERP, 4D-5D, LLM mašīnmācīšanās un mākslīgais intelekts, kā arī raksti, kas publicēti Eiropas būvniecības nozares publikācijās. Viens no maniem ievērojamiem sasniegumiem ir "BIM vēstures" izveide. [111], kas ir visaptveroša karte ar svarīgiem programmatūras risinājumiem datu pārvaldībai būvniecības nozarē. Mana 7 rakstu sērija "BIM attīstības un lobēšanas spēles", kas tulkota vairākās valodās, ir guvusi plašu atzinību kā mēģinājums izgaismot digitālo standartu slēpto dinamiku.

Tā es pārgāju no iežu ieguves uz būvniecības datu ieguvi un sistematizēšanu. Es vienmēr esmu atvērts profesionālam dialogam, jaunām idejām un kopīgiem projektiem. Es ar pateicību pieņemšu jebkuru atsauksmi un gaidu jūsu ziņojumus vai redzēšu jūs starp saviem sociālo tīklu sekotājiem. Liels paldies, ka izlasījāt šo grāmatu līdz galam! Es būšu priecīgs, ja šī grāmata palīdzēs jums labāk izprast tēmu par datiem būvniecības nozarē.

## **TURPMĀKĀS ATTIECĪBAS**

Lasītāju viedokļiem ir liela nozīme turpmākajā publikāciju izstrādē un prioritāro tēmu izvēlē. Īpaši vērtīgi ir komentāri par to, kuras idejas ir izrādījušās noderīgas un kuras ir radījušas šaubas un prasa papildu skaidrojumus vai avotu citēšanu. Grāmatā iekļauts plašs materiālu un analīzes klāsts, no kuriem daži var šķist pretrunīgi vai subjektīvi. Ja lasīšanas gaitā atklāsiet neprecizitātes, nepareizi citētus avotus, loģiskas pretrunas vai pārrakstīšanās klūdas, būšu pateicīgs par jūsu komentāriem, pārdomām vai kritiku, ko varat sūtīt uz šādu adresi:[boikoartem@gmail.com](mailto:boikoartem@gmail.com). Vai izmantojot ziņojumus LinkedIn:  
[linkedin.com/in/boikoartem](https://linkedin.com/in/boikoartem)

Es būšu pateicīgs, ja sociālajos medijos tiks pieminēta grāmata "Uz datiem balstīta būvniecība" - dalīšanās ar lasīšanas pieredzi palīdz izplatīt informāciju par atvērtajiem datiem un rīkiem un atbalsta manu darbu.

## **TULKOJUMA KOMENTĀRS**

Šī grāmata tika tulkota, izmantojot mākslīgā intelekta tehnoloģiju. Tas tulkošanas procesu padarīja daudz ātrāku. Tomēr, kā jebkurā tehnoloģiskā operācijā, var gadīties klūdas vai neprecizitātes. Ja pamanāt kaut ko, kas šķiet nepareizi vai nepareizi iztulkots, lūdzu, rakstiet man. Jūsu komentāri palīdzēs uzlabot tulkojuma kvalitāti.

## **UZ DATIEM BALSTĪTAS BŪVNIECĪBAS KOPIENAS**

Tā ir vieta, kur varat brīvi uzdot jautājumus un dalīties ar savām problēmām un risinājumiem:

DataDrivenConstruction.io: <https://datadrivenconstruction.io>

LinkedIn: <https://www.linkedin.com/company/datadrivenconstruction/>

Twitter: <https://twitter.com/datadrivenconst>

Telegram: <https://t.me/datadrivenconstruction>

YouTube: <https://www.youtube.com/@datadrivenconstruction>

## CITAS PRASMES UN JĒDZIENI

Papildus galvenajiem darba ar datiem būvniecības nozarē principiem DataDrivenConstruction aplūko plašu papildu koncepciju, programmu un prasmju klāstu, kas ir būtiski profesionālim, kurš strādā ar datiem. Dažas no tām ir izklāstītas tikai pārskata veidā, taču tās ir ļoti svarīgas praksē.

Ieinteresētais lasītājs var apmeklēt DataDrivenConstruction.io tīmekļa vietni, kur atrodamas saites uz papildu materiāliem par galvenajām prasmēm. Šie materiāli ietver darbu ar Python un Pandas, ETL - procesu veidošanu, datu apstrādes piemērus būvniecības CAD projektos, lielo datu sistēmas un modernas pieejas būvniecības datu vizualizācijai un analīzei.

Sagatavojot grāmatu "DataDrivenConstruction" un visus gadījumu pētījumus, tika izmantoti daudzi atvērtā koda rīki un programmatūra. Autors vēlas pateikties šādu risinājumu izstrādātājiem un līdzautoriem:

- Python un Pandas - datu apstrādes un automatizācijas pamats
- Scipy, NumPy, Matplotlib un Scikit-Learn - datu analīzes un mašīnmācīšanās bibliotēkas.
- SQL un Apache Parquet - rīki liela apjoma būvniecības datu glabāšanai un apstrādei.
- Atvērtā koda CAD (BIM) atvērto datu rīki atvērtos formātos
- N8n, Apache Airflow, Apache NiFi - orķestrēšanas un darba plūsmas automatizācijas sistēmas.
- DeepSeek, LlaMa, Mistral - Open Source LLM

Īpašs paldies visiem diskusiju dalībniekiem par atvērto datu un rīku tēmu profesionālajās kopienās un sociālajos tīklos, kuru izteiktā kritika, komentāri un idejas palīdzēja uzlabot šīs grāmatas saturu un struktūru.

Sekojiet līdzi projekta attīstībai DataDrivenConstruction.io vietnē, kur tiek publicēti ne tikai grāmatas atjauninājumi un labojumi, bet arī jaunas nodalas, pamācības un praktiski piemēri aprakstīto metožu pielietošanai.

## MAKSIMĀLI ĒRTA DRUKĀTĀ VERSIJA

Jūsu rīcībā ir bezmaksas digitālā versija **Data-Driven Construction**. Ērtākam darbam un ātrākai piekļuvei materiāliem iesakām pievērst uzmanību **drukātajam izdevumam**:



■ **Vienmēr pa rokai:** grāmata drukātā formātā klūs par uzticamu darba rīku, kas ļaus ātri atrast un izmantot vajadzīgās vizualizācijas un diagrammas jebkurā darba situācijā.

■ **Augstas kvalitātes ilustrācijas:** visi attēli un grafikas drukātajā izdevumā ir maksimāli kvalitatīvi.

■ **Ātra piekļuve informācijai:** vienkārša navigācija, iespēja veikt piezīmes, grāmatzīmes un strādāt ar grāmatu jebkurā vietā.

legādājoties pilnu drukāto grāmatas versiju, jūs iegūstat ērtu rīku ērtam un efektīvam darbam ar informāciju: iespēju ātri izmantot vizuālos materiālus ikdienas uzdevumos, ātri atrast nepieciešamās shēmas un veikt piezīmes. Turklāt ar savu pirkumu atbalstāt atvērto zināšanu izplatīšanu.

Grāmatas drukāto versiju var pasūtīt: [datadrivenconstruction.io/books](http://datadrivenconstruction.io/books)



## UNIKĀLA IESPĒJA STRATĒGISKĀ POZICIONĒŠANA

Mēs piedāvājam jums izvietot reklāmas materiālus DataDrivenConstruction bezmaksas versijā. Izdevuma maksas versija jau pirmajā gadā pēc tās iznākšanas ir piesaistījusi speciālistu uzmanību vairāk nekā 50 pasaules valstīs - no Latīnamerikas līdz Āzijas un Klusā okeāna reģionam. Lai apspriestu individuālus sadarbības nosacījumus un saņemtu detalizētu informāciju par izvietošanas iespējām, lūdzu, aizpildiet atsauksmju veidlapu oficiālajā portālā [datadrivenconstruction.io](http://datadrivenconstruction.io) vai rakstiet grāmatas beigās norādītajiem kontaktiem.



## GRĀMATAS NODAĻAS IR PIEEJAMAS VIETNĒ DATADRIVENCONSTRUCTION.IO

Datu vadītas būvniecības nodaļas varat lasīt tīmekļa vietnē Data-Driven Construction, kur pakāpeniski tiek publicētas grāmatas sadaļas, lai jūs varētu ātri atrast nepieciešamo informāciju un izmantot to savā darbā. Tur atradīsiet arī daudzas citas publikācijas par līdzīgām tēmām, kā arī lietojumu un risinājumu piemērus, kas palīdzēs jums attīstīt savas prasmes un pielietot datus būvniecībā.



## JAUNĀKĀS GRĀMATAS VERSIJAS LEJUPIELĀDĒT NO OFICIĀLĀS TĪMEKLĀ VIETNES

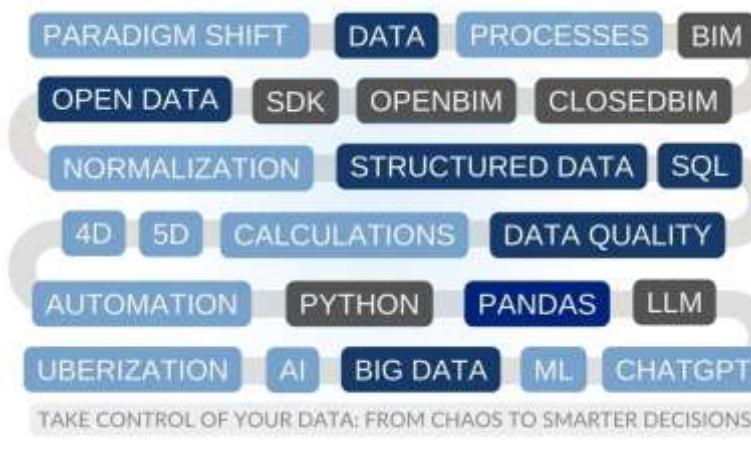
Pašreizējās un jaunākās DataDrivenConstruction grāmatas versijas ir pieejamas lejupielādei vietnē [datadrivenconstruction.io](http://datadrivenconstruction.io). Ja vēlaties saņemt jaunumus ar jaunām grāmatas nodaļām, praktiskiem padomiem vai pārskatu par jaunām lietojumprogrammām, abonējiet jaunumus:

- Jūs būsiet pirmais, kas tiks iepazīstināts ar jaunajām grāmatas sadaļām.
- Saņemiet praktisku gadījumu izpēti un padomus par analītiku un automatizāciju būvniecības nozarē.
- Sekojiet līdzi tendencēm, publikācijām un lietojumu piemēriem.

Dodies uz [datadrivenconstruction.io](http://datadrivenconstruction.io), lai parakstītos!

# UZ DATIEM BALSTĪTA KONSTRUKCIJA: KONSULTĀCIJAS, SEMINĀRI UN APMĀCĪBAS

DataDrivenConstruction mācību programmas un konsultācijas ir palīdzējušas desmitiem vadošo būvniecības uzņēmumu visā pasaulei palielināt efektivitāti, samazināt izmaksas un uzlabot risinājumu kvalitāti. DataDrivenConstruction klientu vidū ir daži no lielākajiem spēlētājiem miljardu eiro vērtajā tirgū, tostarp būvniecības, konsultāciju un IT uzņēmumi.



Kāpēc izvēlēties mūs?

- **Aktualitāte:** sarunas par nozares galvenajām tendencēm un atziņām.
- **Praktiski:** palīdzēt profesionāļiem efektīvi risināt ikdienas uzdevumus, izmantojot PoC.
- **Individuāla pieeja:** pielāgota jūsu uzņēmumam, maksimāli palielinot mācību un konsultāciju priekšrocības.

DataDrivenConstruction komandas galvenās darbības jomas:

- **Datu kvalitātes pārvaldība:** palīdz parametrizēt uzdevumus, apkopot prasības, validēt un sagatavot datus automatizētai apstrādei.
- **Datu ieguve - datu ieguve un strukturēšana:** izveidojiet ETL procesus un iegūstiet datus no e-pasta ziņojumiem, PDF, Excel, attēliem un citiem avotiem.
- **BIM un CAD analītika:** vāc, strukturē un analizē informāciju no RVT failiem, IFC, DWG un citiem CAD (BIM) formātiem.
- **Analītika un datu pārveidošana:** atšķirīgas informācijas pārvēršana strukturētos datos, analīzē, ieskatos un lēmumos.
- **Datu integrācija un procesu automatizācija:** no automatizētas dokumentu izveides līdz integrācijai ar iekšējām sistēmām un ārējām datubāzēm.

Sazinieties ar DataDrivenConstruction.io, lai uzzinātu, kā automatizācijas izmantošana var palīdzēt jūsu uzņēmumam sasniegt taustāmus biznesa rezultātus.

## GLOSSĀRIJS

**Mākslīgais intelekts** - mākslīgais intelekts; datorsistēmu spēja veikt uzdevumus, kuru veikšanai parasti ir nepieciešams cilvēka intelekts, piemēram, modeļu atpazišana, mācīšanās un lēmumu pieņemšana.

**Apache Airflow** ir atvērtā koda darbplūsmu orķestrēšanas platforma, kas ļauj programmatiski izveidot, plānot un izsekot darbplūsmām un ETL, izmantojot DAG (directed acyclic graphs).

**Apache NiFi** ir rīks datu plūsmu automatizēšanai starp sistēmām, kas specializējas datu maršrutēšanā un pārveidošanā.

**Apache Parquet** ir efektīvs failu formāts kolonnāriem datu uzglabāšanas veidiem, kas optimizēts izmantošanai lielo datu analīzes sistēmās. Tas nodrošina ievērojamu saspiešanu un ātru apstrādi.

**API (lietojumprogrammu saskarne)** - formalizēta saskarne, kas ļauj vienai programmai mijiedarboties ar citu programmu bez piekļuves pirmkodam, apmainoties ar datiem un funkcijām, izmantojot standartizētus pieprasījumus un atbildes.

**Atribūts** - objekta īpašība, kas raksturo tā iezīmes (piemēram, platību, tilpumu, izmaksas, materiālu).

**Datu bāzes** ir organizētas struktūras informācijas glabāšanai, pārvaldībai un piekļuvei, ko izmanto efektīvai datu izgūšanai un apstrādei.

**BEP (BIM izpildes plāns)** - ēkas informācijas modelēšanas ieviešanas plāns, kas nosaka mērķus, metodes un procesus BIM ieviešanai projektā.

**Lielie dati** - liela apjoma, daudzveidības un atjaunināšanas ātruma informācijas masīvi, kuru apstrādei un analīzei nepieciešamas īpašas tehnoloģijas.

**BI (Business Intelligence) - Biznesa izlūkošana;** procesi, tehnoloģijas un rīki, lai pārveidotu datus jēgpilnā informācijā lēmumu pieņemšanai.

**BIM (Building Information Modeling)** - Būvniecības **informācijas** modelēšana; būvniecības projektu fizisko un funkcionālo īpašību digitālās reprezentācijas radīšanas un pārvaldības process, kas ietver ne tikai 3D modeļus, bet arī informāciju par veikspēju, materiāliem, termiņiem un izmaksām.

**BlackBox/WhiteBox** - pieejas sistēmas izpratnei: pirmajā gadījumā iekšējā logika ir slēpta, redzamas tikai ieejas un izejas; otrajā gadījumā apstrāde ir pārredzama un pieejama analīzei.

**Bounding Box** ir ģeometriskā konstrukcija, kas apraksta objekta robežas trīsdimensiju telpā, izmantojot minimālās un maksimālās X, Y un Z koordinātas, veidojot "kastīti" ap objektu.

**BREP (Boundary Representation)** ir objektu ģeometriskā attēlošana, kas definē objektus, izmantojot virsmu robežas.

**CAD (Computer-Aided Design)** ir datorizēta projektēšanas sistēma, ko izmanto precīzu rasējumu un 3D modeļu izveidei, redīģēšanai un analīzei arhitektūrā, būvniecībā, inženierzinātnēs un citās nozarēs.

**CAFM (Computer-Aided Facility Management)** ir īpašuma un infrastruktūras pārvaldības programmatūra, kas ietver telpu plānošanu, aktīvu pārvaldību, apkopi un izmaksu uzraudzību.

**CDE (Common Data Environment)** - centralizēta digitālā telpa projekta informācijas pārvaldībai, glabāšanai, koplietošanai un sadarbībai visos objekta dzīves cikla posmos.

**Izcilības centrs (CoE)** ir specializēta struktūra organizācijā, kas atbild par konkrētas jomas zināšanu attīstīšanu, standartu un paraugprakses izstrādi, personāla apmācību un atbalstu inovāciju ieviešanā.

**CoClass** ir mūsdienīga, trešās paaudzes ēku elementu klasifikācijas sistēma.

**Konceptuālais datu modelis** ir augsta līmeņa pamatvienību un to attiecību attēlojums bez atribūtu detaļām, ko izmanto datubāzes projektēšanas sākotnējos posmos.

**CRM (Customer Relationship Management)** ir klientu attiecību pārvaldības sistēma, ko izmanto pārdošanas un apkalpošanas procesu automatizēšanai.

**DAG (Directed Acyclic Graph)** ir virzīts aciklisks grafs, ko izmanto datu orķestrēšanas sistēmās (Airflow, NiFi), lai noteiktu uzdevumu secību un atkarības.

**Dash** ir Python lietojumprogramma interaktīvu tīmekļa datu vizualizāciju izveidei.

**Informācijas paneli - Informācijas panelis**, kas vizuāli attēlo galvenos veikspējas rādītājus un rādītājus reāllaikā.

**Uz datiem orientēta pieeja** ir metodoloģija, kas par prioritāti izvirza datus, nevis lietojumprogrammas vai programmatūras kodu, padarot datus par galveno organizācijas aktīvu.

**Datu pārvaldība** - prakses, procesu un politiku kopums, kas nodrošina atbilstošu un efektīvu datu izmantošanu organizācijā, tostarp piekļuves, kvalitātes un drošības kontroli.

**Datu ezers** ir krātuve, kas paredzēta liela apjoma neapstrādātu datu glabāšanai to sākotnējā formātā, līdz tie tiek izmantoti.

**Datu ezernīca** ir arhitektūras pieeja, kas apvieno datu ezeru (Data Lake) elastību un mērogojamību ar datu noliktavu (DWH) pārvaldāmību un veikspēju.

**Uz datiem balstīta būvniecība** ir stratēģiska pieeja, kurā katru objekta dzīves cikla posmu - no projektēšanas līdz ekspluatācijai - atbalsta automatizētas, savstarpēji saistītas sistēmas. Šī pieeja nodrošina nepārtrauktu, uz faktiem balstītu mācīšanos, samazina nenoteiktību un ļauj uzņēmumiem sasniegt ilgtspējīgu līderpozīciju nozarē.

**Datu integrators** - uzņēmums, kas specializējas datu apvienošanā no dažādiem avotiem un to analīzē, lai pieņemtu vadības lēmumus.

**Uz datiem balstīta pieeja** - metodoloģija, kurā dati tiek uzskatīti par stratēģisku vērtību un lēmumi tiek pieņemti, pamatojoties uz objektīvu informācijas analīzi, nevis subjektīvu viedokli.

**Datu minimālisms** - pieeja, kas ļauj samazināt datus līdz vērtīgākajiem un nozīmīgākajiem, ļaujot vienkāršot informācijas apstrādi un analīzi.

**Datu purvs** - nestukturētu datu izkliedēta masa, kas rodas, ja informācija tiek nekontrolēti vākta un uzglabāta bez pienācīgas organizācijas.

**DataOps** ir metodoloģija, kas apvieno DevOps principus, datus un analītiku, koncentrējoties uz datu plūsmu sadarbības uzlabošanu, integrāciju un automatizāciju.

**Informācijas digitalizācija** ir process, kurā visi būvniecības darbības aspekti tiek pārvērstīgi digitālā formā, kas piemērota analīzei, interpretācijai un automatizācijai.

**DataFrame** - Pandas bibliotēkas divdimensiju tabulārā datu struktūra, kur rindas attēlo atsevišķus ierakstus vai objektus, bet kolonas - to īpašības vai atribūtus.

**Aprakstošā analīze** - vēsturisko datu analīze, lai saprastu, kas noticis pagātnē.

**Diagnostikas analīze** - datu analīze, lai noteiktu, kāpēc kaut kas noticis.

**Ganta diagramma** ir projekta plānošanas rīks, kas attēlo uzdevumus kā horizontālas joslas uz laika līnijas, ļaujot vizualizēt darbu secību un ilgumu.

**DWH (Data Warehouse)** ir centralizēta datu noliktavas sistēma, kas apkopo informāciju no dažādiem avotiem, strukturē to un dara pieejamu analītikai un pārskatu sagatavošanai.

**ESG (Environmental, Social, Governance)** - kritēriju kopums uzņēmuma vai projekta ietekmes uz vidi, sociālo jomu un pārvaldību novērtēšanai.

**ELT (Extract, Load, Transform)** ir process, kurā dati vispirms tiek iegūti no avotiem, ielādēti krātuvē un pēc tam pārveidoti analītiskiem mērķiem.

**ETL (Extract, Transform, Load)** ir datu ieguves process no dažādiem avotiem, to pārveidošana vajadzīgajā formātā un ielādēšana mērķa krātuvē analīzei.

**ER-diagramma (Entity-Relationship)** - vizuāla diagramma, kurā attēlotas datu modelēšanā izmantotās vienības, to atribūti un attiecības starp tām.

**ERP (Enterprise Resource Planning)** ir visaptveroša modulāra uzņēmuma resursu plānošanas sistēma, ko izmanto, lai pārvaldītu un optimizētu dažādus būvniecības procesa aspektus.

**Iezīmes** - mašīnmācībā neatkarīgie mainīgie vai atribūti, ko izmanto kā modeļa ievades datus.

**Fiziskais datu modelis** - detalizēts datu bāzes struktūras attēlojums, tostarp tabulas, kolonas, datu tipi, atslēgas un indeksi, kas optimizēti konkrētai DBVS.

**FPDF** ir Python bibliotēka PDF dokumentu izveidei.

**Geometric Core** ir programmatūras komponents, kas nodrošina pamata algoritmus ģeometrisko objektu izveidei, rediģēšanai un analīzei CAD, BIM un citās inženierprogrammās.

**HiPPO (Highest Paid Person's Opinion)** - pieeja lēmumu pieņemšanai, kas balstīta uz organizācijas visaugstāk atalgošās personas viedokli, nevis objektīviem datiem.

**IDE (integrētā izstrādes vide)** - integrētā izstrādes vide, visaptverošs rīks koda rakstīšanai, testēšanai un atklūdošanai (piemēram, PyCharm, VS Code, Jupyter Notebook).

**IDS (informācijas piegādes specifikācija)** ir informācijas piegādes specifikācija, kas nosaka datu prasības dažādos projekta posmos.

**IFC (Industry Foundation Classes)** ir BIM datu apmaiņas formāts, kas nodrošina savietojamību starp dažādiem programmatūras risinājumiem.

**Rūpniecība 5.0** ir rūpniecības attīstības koncepcija, kas apvieno digitalizācijas, automatizācijas un mākslīgā intelekta iespējas ar cilvēku potenciālu un vides ilgtspējību.

**Datu integrācija** ir process, kurā no dažādiem avotiem iegūtie dati tiek apvienoti vienotā, saskanīgā sistēmā, lai nodrošinātu vienotu priekšstatu par informāciju.

**Informācijas "silosi"** ir izolētas datu glabāšanas sistēmas, kas neveic informācijas apmaiņu ar citām sistēmām, tādējādi radot šķēršļus efektīvai datu izmantošanai.

**IoT (lietiskais internets)** ir koncepcija, kas paredz fizisku objektu savienošanu ar internetu, lai vāktu, apstrādātu un pārsūtītu datus.

**k-NN (k-tuvākie kaimiņi)** ir mašīnmācīšanās algoritms, kas klasificē objektus, pamatojoties uz līdzību ar tuvākajiem kaimiņiem mācību paraugā.

**Kaggle** ir datu analīzes un mašīnmācīšanās sacensību platforma.

**Aprēķins** - būvdarbu vai procesu izmaksu aprēķins noteiktai mērvienībai (piemēram, 1 m<sup>2</sup> ģipškartona sienas, 1 m<sup>3</sup> betona).

**KPI (Key Performance Indicators)** ir galvenie rezultatīvie rādītāji, skaitliski izsakāmi rādītāji, ko izmanto, lai novērtētu uzņēmuma vai konkrēta projekta panākumus.

**Etiketes** - Mašīnmācībā - mērķa mainīgie vai atribūti, kas modelim jāparedz.

**Mācīšanās algoritms** - process, kurā, izmantojot apmācības datu kopumu, tiek atrasta labākā hipotēze modelī, kas atbilst mērķa funkcijai.

**Lineārā regresija** - statistikas metode, ar kuru modelē atkarīgā mainīgā un viena vai vairāku neatkarīgo mainīgo attiecību.

**LLM (Large Language Model) - lielais** valodas modelis, mākslīgais intelekts, kas apmācīts saprast un ģenerēt tekstu no milzīgiem datu apjomiem, spēj analizēt kontekstu un rakstīt programmēšanas kodu.

**LOD (detalizācijas/izstrādes līmenis)** - modeļa detalizācijas līmenis, kas nosaka ģeometriskās precīzitātes un informācijas saturā pakāpi.

**Loģiskais datu modelis** ir detalizēts būtņu, atribūtu, atslēgu un attiecību apraksts, kas atspoguļo uzņēmēdarbības informāciju un noteikumus, un ir starposms starp konceptuālo un fizisko modeli.

**Mašīnmācīšanās** - mākslīgā intelekta metožu klase, kas ļauj datorsistēmām mācīties un prognozēt datus bez tiešas programmēšanas.

**Masterformat** ir pirmās paaudzes klasifikācijas sistēma, ko izmanto, lai strukturētu būvniecības specifikācijas iedalījums un disciplīnās.

**MEP (mehāniskās, elektriskās un santehnikas sistēmas)** - ēku inženiertehniskās sistēmas, kas ietver mehāniskās, elektriskās un santehnikas sastāvdaļas.

**Tīkls** ir trīsdimensiju objektu tīkla veida attēlojums, kas sastāv no virsotnēm, malām un virsmām.

**Modelis** - mašīnmācībā - dažādu hipotēžu kopums, no kurām viena aproksimē prognozējamo vai aproksimējamo mērķa funkciju.

**Datu modelēšana** ir datu un to attiecību strukturētas reprezentācijas izveides process, kas paredzēts ieviešanai informācijas sistēmās, ieskaitot konceptuālo, loģisko un fizisko līmeni.

**n8n** ir atklātā pirmkoda rīks, kas ļauj automatizēt darbplūsmas un integrēt lietojumprogrammas, izmantojot zema kodēšanas līmeņa pieejumu.

**Normalizēšana** - mašīnmācīšanās procesā - dažādu skaitlisko datu pielīdzināšana vienotai skalai, lai atvieglotu apstrādi un analīzi.

**Reversā inženierija** - priekšmeta ierīces, darbības un ražošanas tehnoloģijas izpētes process, analizējot tā struktūru, funkcijas un darbību. Datu kontekstā - informācijas iegūšana no patentētiem formātiem izmantošanai atvērtās sistēmās.

**OCR (optiskā rakstzīmu atpazīšana)** ir optiskā rakstzīmu atpazīšanas tehnoloģija, kas pārvērš teksta attēlus (skenētus dokumentus, fotogrāfijas) mašīnlasāmā teksta formātā.

**OmniClass** ir otrs paaudzes starptautisks klasifikācijas standarts būvniecības informācijas pārvaldībai.

**Ontoloģija** - jēdzienu savstarpējo attiecību sistēma, kas formalizē konkrētu zināšanu jomu.

**Atvērtā pirmkoda programmatūra** - atvērtā pirmkoda programmatūras izstrādes un izplatīšanas modelis, kas ir pieejams brīvai lietošanai, izpētei un modificēšanai.

**Open BIM** ir atvērta BIM koncepcija, kas ietver atvērtu standartu un formātu izmantošanu datu apmaiņai starp dažādiem programmatūras risinājumiem.

**Atvētie standarti** - publiski pieejamas specifikācijas konkrēta uzdevuma izpildei, kas ļauj dažādām sistēmām mijiedarboties un apmainīties ar datiem.

**Pandas** ir atvērtā pirmkoda Python bibliotēka datu apstrādei un analīzei, kas nodrošina DataFrame un Series datu struktūras, lai efektīvi apstrādātu tabulāro informāciju.

**Atvērto datu paradigma** ir pieeja datu apstrādei, kurā informācija ir brīvi pieejama izmantošanai, atkārtotai izmantošanai un izplatīšanai ikvienam.

**Parametriskā** metode ir būvprojekta novērtēšanas metode, kurā izmanto statistiskos modeļus, lai novērtētu vērtību, pamatojoties uz projekta parametriem.

**PIMS (projekta informācijas modelis)** ir digitāla sistēma, kas paredzēta visas projekta informācijas organizēšanai, glabāšanai un koplietošanai.

**Caurulvads** - datu apstrādes procesu secība, sākot no ieguves un pārveidošanas līdz analīzei un vizualizācijai.

**PMIS (Project Information Management System)** ir projektu vadības sistēma, kas paredzēta detalizētai uzdevumu kontrolei atsevišķa būvniecības projekta līmenī.

**Prognozējošā** analītika ir analītikas sadaļa, kurā izmanto statistikas metodes un mašīnmācīšanos, lai prognozētu nākotnes rezultātus, pamatojoties uz vēsturiskiem datiem.

**Preskriptīvā** analītika - analītikas sadaļa, kas ne tikai prognozē nākotnes rezultātus, bet arī iesaka optimālas darbības, lai sasniegtu vēlamos rezultātus.

**Patentēti** formāti ir slēgti datu formāti, ko kontrolē konkrēts uzņēmums, kas ierobežo informācijas apmaiņas iespējas un palielina atkarību no konkrētas programmatūras.

**QTO (Quantity Take-Off)** ir process, kurā no projekta dokumentiem tiek iegūti elementu daudzumi, lai aprēķinātu projekta īstenošanai nepieciešamo materiālu daudzumus.

**Kvalitātes vadības sistēma** - kvalitātes vadības sistēma, kas nodrošina procesu un rezultātu atbilstību noteiktajām prasībām.

**RAG (Retrieval-Augmented Generation)** ir metode, kas apvieno valodas modeļu ģeneratīvās iespējas ar attiecīgās informācijas ieguvi no uzņēmumu datubāzēm, uzlabojot atbilžu precizitāti un atbilstību.

**RDBMS (Relational Database Management System)** ir relāciju datubāzes pārvaldības sistēma, kas organizē informāciju savstarpēji saistītu tabulu veidā.

**RegEx (regulārās izteiksmes)** ir formalizēta virķu meklēšanas un apstrādes valoda, kas ļauj norādīt veidnes teksta datu pārbaudei, lai pārbaudītu to atbilstību noteikiem kritērijiem.

**Regresija** ir statistikas metode, ar kuru analizē attiecības starp mainīgajiem lielumiem.

**CO<sub>2</sub> aprēķini** ir oglekļa dioksīda emisiju, kas saistītas ar būvmateriālu un procesu ražošanu un izmantošanu, novērtēšanas metode.

**Resursu metode** ir aprēķinu metode, kuras pamatā ir visu būvdarbu veikšanai nepieciešamo resursu (materiālu, darbaspēku, iekārtu) detalizēta analīze.

**RFID (radiofrekvenču identifikācija)** ir tehnoloģija automātiskai objektu identificēšanai, izmantojot radio signālus, ko izmanto materiālu, iekārtu un personāla izsekošanai.

**ROI (Return on Investment)** ir rādītājs, kas atspoguļo attiecību starp peļņu un ieguldītajiem līdzekļiem un ko izmanto, lai novērtētu ieguldījumu efektivitāti.

**SaaS (programmatūra kā pakalpojums)** ir programmatūras kā pakalpojuma modelis, kurā lietojumprogrammas mitina pakalpojumu sniedzējs un tās ir pieejamas lietotājiem internetā.

**SCM (Supply Chain Management)** - piegādes ķēdes pārvaldība, kas ietver visu procesu koordināciju un optimizāciju, sākot no materiālu iegādes līdz gatavās produkcijas piegādei.

**Datu silosi** ir izolēti informācijas krājumi organizācijā, kas nav integrēti ar citām sistēmām, tādējādi apgrūtinot datu apmaiņu un padarot to neefektīvu.

**SQL** (strukturētā vaicājumu **valoda**) ir strukturēta vaicājumu valoda, ko izmanto darbā ar relāciju datubāzēm.

**SQLite** ir viegla, iebūvējama, daudzplatformu DBVS, kurai nav nepieciešams atsevišķs serveris un kura atbalsta SQL pamatfunkcijas, ko plaši izmanto mobilajās lietojumprogrammās un iebūvētajās sistēmās.

**Strukturēti dati** - informācija, kas sakārtota noteiktā formātā ar skaidru struktūru, piemēram, relāciju datubāzēs vai tabulās.

**Vāji strukturēti dati** - informācija ar daļēju organizāciju un elastīgu struktūru, piemēram, JSON vai XML, kur dažādi elementi var saturēt dažādus atribūtu kopumus.

Entitāte ir konkrēts vai abstrakts reālās pasaules objekts, ko var unikāli identificēt, aprakstīt un attēlot datu veidā.

**Uzraudzīta mācīšanās** - mašīnmācīšanās veids, kurā algoritms tiek apmācīts, izmantojot sadalītus datus, kad katram piemēram ir zināms vēlamais rezultāts.

**Taksonomija** ir hierarhiska klasifikācijas sistēma, ko izmanto, lai sistemātiski kategorizētu elementus, pamatojoties uz kopīgām pazīmēm.

**Titanic datu kopa** ir populāra datu kopa mašīnmācīšanās modeļu apmācībai un testēšanai.

**Apmācība** - process, kurā mašīnmācīšanās algoritms analizē datus, lai noteiku modeļus un izveidotu modeli.

**Pārneses mācīšanās** ir mašīnmācīšanās metode, kurā vienam uzdevumam apmācīts modelis tiek izmantots kā sākumpunkts citam uzdevumam.

**Transformācija** - datu formāta, struktūras vai satura maiņas process, lai tos vēlāk varētu izmantot.

**Datu prasības** - formalizēti kritēriji, kas nosaka uzņēmējdarbības procesu atbalstam nepieciešamās informācijas struktūru, formātu, pilnīgumu un kvalitāti.

**Būvniecības nozares überizācija** ir tradicionālo būvniecības uzņēmējdarbības modeļu transformācijas process digitālo platformu ietekmē, kas nodrošina tiešu mijiedarbību starp klientiem un darbuzņēmējiem bez starpniekiem.

**Uniclass** ir otrās un trešās paaudzes ēku elementu klasifikācijas sistēma, ko plaši izmanto Apvienotajā Karalistē.

**USD (Universal Scene Description)** ir datu formāts, kas izstrādāts datorgrafikas vajadzībām, bet ir ieguvis pielietojumu inženiertehniskajās sistēmās, pateicoties tā vienkāršai struktūrai un neatkarībai no geometriskiem kodoliem.

**Datu validēšana** ir informācijas pārbaudes process, kurā tiek pārbaudīta atbilstība noteiktiem kritērijiem un prasībām, lai nodrošinātu datu precīzitāti, pilnīgumu un konsekvenči.

**Vektoru** datubāze - specializēta tipa datubāze, kurā dati tiek glabāti kā daudzdimensiju vektori, lai nodrošinātu efektīvu semantisko meklēšanu un objektu salīdzināšanu.

**Vektoru attēlošana (iestrādāšana)** ir metode, ar kuras palīdzību datus pārveido daudzdimensiju skaitliskos vektoros, kas ļauj mašīniskiem algoritmiem efektīvi apstrādāt un analizēt informāciju.

**VectorOps** ir metodoloģija, kas vērsta uz daudzdimensiju vektoru datu apstrādi, glabāšanu un analīzi, kas īpaši svarīga tādās jomās kā digitālie dvīņi un semantiskā meklēšana.

**Vizualizācija** - datu grafiska attēlošana, lai labāk uztvertu un analizētu informāciju.

Termini tika iedalīti alfabētiskā kategorijā pēc to nosaukumiem angļu valodā.

## ATSAUČU UN TIEŠSAISTES MATERIĀLU SARAKSTS

- [1] Gartner, "IT Key Metrics Data 2017: Index of Public Documents and Metrics", 2016. gada 12. decembris. [Online]. Pieejams: <https://www.gartner.com/en/documents/3530919>. [Piekļuves datums: 2025. gada 1. marts].
- [2] KPMG, "Pazīstami izaicinājumi - jaunas pieejas. 2023. gada pasaules būvniecības apsekojums", 2023. gada 1. janvāris. [Online]. Pieejams: <https://assets.kpmgsites.com/dam/kpmgsites/xx/pdf/2023/06/familiar-challenges-new-solutions-1.pdf>. [Piekļuve: 2025. gada 5. marts].
- [3] F. R. Barnards, "Attēls ir tūkstoš vārdu vērts", 10 Mapi 1927. [Online]. Pieejams: [https://en.wikipedia.org/wiki/A\\_picture\\_is\\_worth\\_a\\_thousand\\_words](https://en.wikipedia.org/wiki/A_picture_is_worth_a_thousand_words). [Atsauces datums: 2025. gada 15. marts].
- [4] M. Bastian, "Microsoft izpilddirektors Satya Nadella saka, ka AGI ir "bezjēdzīga etalona uzlaušana", 21 Feb. 2025. [Online]. Pieejams: <https://the-decoder.com/microsoft-ceo-satya-nadella-says-self-claimingagi-is-nonsensical-benchmark-hacking/>. [Piekļuves datums: 2025. gada 15. marts].
- [5] W. E. Forums "Forums, kas veido būvniecības nākotni - ainava pārmaiņu procesā", 2016. gada 1. janvāris. [Online]. Pieejams: [https://www3.weforum.org/docs/WEF\\_Shaping\\_the\\_Future\\_of\\_Construction.pdf](https://www3.weforum.org/docs/WEF_Shaping_the_Future_of_Construction.pdf). [Piekļuves datums: 2025. gada 2. marts].
- [6] C. D. Gillespie, "Māls: Zemes iejaukšanās māla laikmetā", 2024. gads. [Online]. Pieejams: <https://ufl.pb.unizin.org/imos/chapter/clay/>.
- [7] "3. gadsimta p.m.ē. papiruss Valoda ir grieķu," 2024. [Online]. Available: <https://www.facebook.com/429710190886668/posts/595698270954525>.
- [8] "Monitorings: pieejamo rīku izmantošana", 1980. gads. [Online]. Pieejams: <https://pubmed.ncbi.nlm.nih.gov/10246720/>. [Uzrunas datums: 2025. gada 15. marts].
- [9] PWC, "Dati, kas skolēniem nepieciešami, lai gūtu panākumus strauji mainīgajā uzņēmējdarbības pasaulē", 2015. gada 15. februāris. [Online]. Pieejams: <https://www.pwc.com/us/en/faculty-resource/assets/PwC-Data-driven-paper-Feb2015.pdf>. [Piekļuves datums: 2025. gada 15. marts].
- [10] Skanska ASV, "Rudens būvniecības tirgus tendences", 2023. gada 2. novembris. [Online]. Available: <https://x.com/SkanskaUSA/status/1720167220817588714>.

- [11] "Oxford Essential Quotations (4th ed.)", Oxford University Press, 2016. [Online]. Available: <https://www.oxfordreference.com/display/10.1093/acref/9780191826719.001.0001/q-or-ed4-00006236>. [Piekļuves datums: 2025. gada 1. marts].
- [12] "Citāts: Sondergārds par datu analītiku," [Online]. Pieejams: <https://www.causeweb.org/cause/resources/library/r2493>. [Piekļuves datums: 2025. gada 15. marts].
- [13] "Kā globālā interese par mākslīgo intelektu palielina datu pārvaldības tirgu", 28 May 2024. [Online]. Pieejams: <https://iot-analytics.com/how-global-ai-interest-is-boosting-data-management-market/>. [Piekļuves datums: 2025. gada 15. marts].
- [14] И. McCue, "ERP vēsture", 2024. [Online]. Pieejams: <https://www.netsuite.com/portal/resource/articles/erp/erp-history.shtml>.
- [15] erpscout, "ERP cena: cik maksā ERP sistēma?", [Online]. Pieejams: <https://erpscout.de/en/erp-costs/>. [Piekļuves datums: 2025. gada 15. marts].
- [16] softwarepath, "Ko 1384 ERP projekti mums stāsta par ERP izvēli (2022. gada ERP ziņojums)", 2022. gada 18. janvāris. [Online]. Pieejams: <https://softwarepath.com/guides/erp-report>. [Piekļuves datums: 2025. gada 15. marts].
- [17] Deloitte, "Uz datiem balstīta pārvaldība digitālajos kapitāla projektos", 2016. gada 16. decembris. [Online]. Pieejams: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/Real%20Estate/us-engineering-construction-data-driven-management-digital-capital-projects.pdf>. [Adreses datums: 2025. gada 1. marts].
- [18] McKinsey, "Uz datiem balstīts uzņēmums 2025. gadā", 2022. gada 28. janvāris. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-data-driven-enterprise-of-2025>. [Atsauces datums: 2024. gada 22. maijs].
- [19] Wikipedia, "Moore's law," [Online]. Pieejams: [https://en.wikipedia.org/wiki/Moore%27s\\_law](https://en.wikipedia.org/wiki/Moore%27s_law). [Piekļuves datums: 2025. gada 15. marts].
- [20] Accenture, "Lielākas vērtības radīšana ar kapitāla projektiem", 2020. gada 1. janvāris. [Online]. Pieejams: <https://www.accenture.com/content/dam/accenture/final/a-com-migration/r3-3/pdf/pdf-143/accenture-industryx-building-value-capital-projects-highres.pdf>. [Piekļuves datums: 2024. gada 3. marts].
- [21] B. Marr, "Cik daudz datu mēs radām katru dienu? Prātam neaptverami statistikas dati, kas būtu jāizlasa ikviename," 2018. gads. [Online]. Pieejams: <https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-create-every-day-the-mind-blowing-stats-everyone-should-read>.

- [22] "Cik daudz datu tiek radīts katru dienu?", 2024. gads. [Online]. Pieejams: <https://graduate.northeastern.edu/resources/how-much-data-produced-every-day/>.
- [23] T. Sullivan, "Mākslīgais intelekts un globālā "datasfēra": cik daudz informācijas būs cilvēci līdz 2025. gadam?", 2024. gads. [Online]. Pieejams: <https://www.datauniverseevent.com/en-us/blog/general/AI-and-the-Global-Datasphere-How-Much-Information-Will-Humanity-Have-By-2025.html>.
- [24] Statista, "Kopējais drukāto grāmatu skaits, kas izdotas dažādos Rietumeiropas reģionos katrā pusgadsimtā no 1454. līdz 1800. gadam", [Internets]. [Online]. Pieejams: <https://www.statista.com/statistics/1396121/europe-book-production-half-century-region-historical/>. [Atsauces datums: 2025. gada 1. marts].
- [25] "Cenu veidošanas piemēri", 2024. [Online]. Pieejams: <https://cloud.google.com/storage/pricing-examples>.
- [26] M. Ashare, "Uzņēmumi nodod datu glabāšanu ārpakalpojumu sniedzējiem, jo sarežģītība pieauga," 10 May 2024. [Online]. Pieejams: <https://www.ciodive.com/news/enterprises-outsource-data-storage-complexity-rises/715854/>. [Piekļuves datums: 2025. gada 15. marts].
- [27] JETSOFTPRO, "SaaS ir miris? Microsoft izpilddirektora šokējošās prognozes skaidrojums," 13 Jan. 2025. gada 13. janvāris. [Online]. Pieejams: <https://jetsoftpro.com/blog/saas-is-dead/>.
- [28] BG2 Pod, "Satya Nadella | BG2 w/ Bill Gurley & Brad Gerstner", 2024. gada 12. decembris. [Online]. Pieejams: [https://www.youtube.com/watch?v=9NtsnzRFJ\\_o](https://www.youtube.com/watch?v=9NtsnzRFJ_o). [Piekļuves datums: 2025. gada 15. marts].
- [29] GoodReads, "Tim Berners-Lee," [Online]. Pieejams: <https://www.goodreads.com/quotes/8644920-data-is-a-precious-thing-and-will-last-longer-than>. [Piekļuves datums: 2025. gada 15. marts].
- [30] KPMG, "Cue Construction 4.0: Make-or-Break Time", 2023. gada 1. janvāris. [Online]. Pieejams: <https://kpmg.com/ca/en/home/insights/2023/05/cue-construction-make-or-break-time.html>. [Piekļuve: 2025. gada 5. marts].
- [31] I. Deininger, B. Koch, R. Bauknecht, and M. Langhans, "Using digital models for decarbonisation of a production site: An example of connecting a building model, a production model and an energy model," 2024. [Online]. Pieejams: [https://www.researchgate.net/publication/374023998\\_Using\\_Digital\\_Models\\_to\\_Decarbonize\\_a\\_Production\\_Site\\_A\\_Case\\_Study\\_of\\_Connecting\\_the\\_Building\\_Model\\_Production\\_Model\\_and\\_Energy\\_Model](https://www.researchgate.net/publication/374023998_Using_Digital_Models_to_Decarbonize_a_Production_Site_A_Case_Study_of_Connecting_the_Building_Model_Production_Model_and_Energy_Model).
- [32] McKinsey, "REINVENTING CONSTRUCTION: A ROUTE TO HIGHER PRODUCTIVITY", 2017. gada 1. februāris. [Online]. Pieejams:

- <https://www.mckinsey.com/~media/mckinsey/business%20functions/operations/our%20insights/reinventing%20construction%20through%20a%20productivity%20revolution/mgi-reinventing-construction-a-route-to-higher-productivity-full-report.pdf>.
- [33] Būvniecības darba grupa premjerministra vietniekam, "Rethinking Construction", 2014. gada 1. oktobris. [Online]. Pieejams: [https://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking\\_construction\\_report.pdf](https://constructingexcellence.org.uk/wp-content/uploads/2014/10/rethinking_construction_report.pdf).
- [34] Forbes, "Bez viedokļa jūs esat tikai vēl viens cilvēks ar datiem," 2016. gada 15. marts. [Tiešsaistē]. Pieejams: <https://www.forbes.com/sites/silberzahnjones/2016/03/15/without-an-opinion-youre-just-another-person-with-data/>. [Piekļuves datums: 2025. gada 15. marts].
- [35] Wikiquote, "Charles Babbage," [Online]. Pieejams: [https://en.wikiquote.org/wiki/Charles\\_Babbage](https://en.wikiquote.org/wiki/Charles_Babbage). [Piekļuves datums: 2025. gada 15. marts].
- [36] SAP, "Jaunā pētījumā konstatēts, ka gandrīz puse vadītāju uzticas mākslīgajam intelektam vairāk nekā paši sev," 2025. gada 12. marts. [Online]. Pieejams: <https://news.sap.com/2025/03/new-research-executive-trust-ai/>. [Piekļuves datums: 2025. gada 15. marts].
- [37] Kanādas Būvniecības asociācija un KPMG Kanādā, 2021, "Būvniecība digitālajā pasaule", 2021. gada 1. maijs. [Online]. Pieejams: <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [Piekļuves datums: 2025. gada 5. marts].
- [38] ZCS, "Piektās industriālās revolūcijas dekodēšana". [Online]. Pieejams: <https://www.pwc.in/decoding-the-fifth-industrial-revolution.html>. [Piekļuves datums: 2025. gada 15. marts].
- [39] M. K. Privātās tiesības un sabiedriskās problēmas: globālā ekonomika, Peterson Inst. for Intern. Economics, 2012.
- [40] F. N. a. Y. Z. Hārvarda Biznesa skola: Manuel Hoffmann, "Atvērtā pirmkoda programmatūras vērtība", 2024. gada 24. janvāris. [Online]. Pieejams: <https://www.hbs.edu/faculty/Pages/item.aspx?num=65230>. [Piekļuves datums: 2025. gada 15. marts].
- [41] Naval Centre for Cost Analysis Air Force Cost Analysis Agency, "Software Development Cost Estimating Handbook", 2008. gada 1. septembris. [Online]. Pieejams: <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/SW%20Cost%20Est%20Manual%20Vol%20I%20rev%2010.pdf>.
- [42] McKinsey, "Celtniecības produktivitātes uzlabošana," [Online]. Pieejams: <https://www.mckinsey.com/capabilities/operations/our-insights/improving-construction->

- productivity. [Adreses datums: 2025. gada 15. marts].
- [43] A. G. a. C. Syverson, "The Strange and Awful Path of Productivity in the US Construction Sector," 19 Jan. 2023. [Online]. Pieejams: <https://bfi.uchicago.edu/insight/research-summary/the-strange-and-awful-path-of-productivity-in-the-us-construction-sector/>. [Piekļuves datums: 2025. gada 1. marts].
- [44] McKinsey, "Celtniecības produktivitātes nodrošināšana vairs nav obligāta", 2024. gada 9. augusts. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/operations/our-insights/delivering-on-construction-productivity-is-no-longer-optional>. [Piekļuves datums: 2025. gada 5. marts].
- [45] ING Group, "Atpalikšana produktivitātes ziņā būvniecībā palielina būvniecības izmaksas", 2022. gada 12. decembris. [Online]. Pieejams: <https://think.ing.com/articles/lagging-productivity-drives-up-building-costs-in-many-eu-countries/>. [Adreses datums: 2025. gada 15. marts].
- [46] M. Bermans, "Microsoft izpilddirektora šokējoša prognoze: "Āgenti aizstās VISU programmatūru", 2024. gada 19. decembris. [Online]. Pieejams: <https://www.youtube.com/watch?v=uGOLYz2pgr8>. [Piekļuves datums: 2025. gada 15. marts].
- [47] Business Insider, "Uzņēmuma Anthropic vadītājs apgalvo, ka pēc 3 līdz 6 mēnešiem mākslīgais intelekts rakstīs 90 % koda, par ko atbildīgi bija programmatūras izstrādātāji," 2025. gada 15. marts. [Online]. Pieejams: <https://www.businessinsider.com/anthropic-ceo-ai-90-percent-code-3-to-6-months-2025-3>. [Piekļuves datums: 2025. gada 30. marts].
- [48] Statista, "Datu bāzu pārvaldības sistēmu (DBVS) popularitātes salīdzinājums pasaulē 2024. gada jūnijā, pēc kategorijas", 2024. gada jūnijs. [Online]. Pieejams: <https://www.statista.com/statistics/1131595/worldwide-popularity-database-management-systems-category/>. [Adreses datums: 2025. gada 15. marts].
- [49] DB-Engines, "DB-Engines reitings", [Online]. Pieejams: <https://db-engines.com/en/ranking>. [Atsauces datums: 2025. gada 15. marts].
- [50] "Stack Overflow izstrādātāju aptauja 2023," 2024. [Online]. Pieejams: <https://survey.stackoverflow.co/2023/>.
- [51] "SQL", 2024. [Tiešsaistē]. Pieejams: <https://en.wikipedia.org/wiki/SQL>.
- [52] "Strukturēti un nestukturēti dati: kāda ir atšķirība?", 2024. gads. [Online]. Pieejams: <https://www.ibm.com/blog/structured-vs-unstructured-data/>.
- [53] DataDrivenConstruction, "Datu formātu salīdzinājums PDF formātā būvniecības projektiem", 23. aprīlis 2024. gada 23. aprīlis. [Online]. Pieejams: <https://datadrivenconstruction.io/wp-content/uploads/2024/10/COMPARISON-OF-DATA-FORMATS-FOR-CONSTRUCTION->

- PROJECTS-1.pdf.
- [54] "Building Information Modeling Whitepaper site", 2003. gads. [Online]. Pieejams: <https://web.archive.org/web/20030711125527/http:/usa.autodesk.com/adsk/servlet/item?id=2255342&siteID=123112>.
  - [55] A. Boyko, "Lobiju kari un BIM attīstība. 5. daļa: BlackRock ir visu tehnoloģiju meistars. Kā korporācijas kontrolē atvērtā pirmkoda kodus," 2024. gads. [Online]. Pieejams: <https://bigdataconstruction.com/autodesk-oracle-blackrock-open-source/>.
  - [56] D. Ušakovs, "Tiešā modelēšana - kam un kāpēc tā vajadzīga? Konkurētspējīgu tehnoloģiju apskats", 2011. gada 14. 11. [Online]. Pieejams: [https://isicad.net/articles.php?article\\_num=14805](https://isicad.net/articles.php?article_num=14805). [Adreses datums: 2025. 02.]
  - [57] C. Eastman un A. Cthers, "Eastman, Charles; And Cthers", 1974. gada septembris. [Online]. Pieejams: <https://files.eric.ed.gov/fulltext/ED113833.pdf>. [Piekļuves datums: 2025. gada 15. marts].
  - [58] D. Ušakovs, "Tiešā modelēšana - kam un kāpēc tā vajadzīga? Konkurējošo tehnoloģiju apskats", 2011. gada 11. novembris. [Online]. Pieejams: [https://isicad.net/articles.php?article\\_num=14805](https://isicad.net/articles.php?article_num=14805). [Atsauces datums: 2025. gada 15. marts].
  - [59] D. Veisbergs, "CAD vēsture", 2022. gada 12. decembris. [Tiešsaistē]. Pieejams: [https://www.shapr3d.com/blog/history-of-cad?utm\\_campaign=cadhISTORYnet](https://www.shapr3d.com/blog/history-of-cad?utm_campaign=cadhISTORYnet). [Atsauces datums: 2025. gada 15. marts].
  - [60] ADSK, "Baltā grāmata par ēku informācijas modelēšanu", 2002. gads. [Online]. Pieejams: [https://web.archive.org/web/20060512180953/http:/images.adsk.com/apac\\_sapac\\_main/files/4525081\\_BIM\\_WP\\_Rev5.pdf#expand](https://web.archive.org/web/20060512180953/http:/images.adsk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand). [Atsauces datums: 2025. gada 15. marts].
  - [61] ADSK, "Baltā grāmata "Building Information Modeling in Practice"," [Online]. Pieejams: [https://web.archive.org/web/20060512181000/http:/images.adsk.com/apac\\_sapac\\_main/files/4525077\\_BIM\\_in\\_Practice.pdf](https://web.archive.org/web/20060512181000/http:/images.adsk.com/apac_sapac_main/files/4525077_BIM_in_Practice.pdf). [Piekļuves datums: 2025. gada 15. marts].
  - [62] A. Boyko, "Lobiju kari un BIM attīstība. 2. daļa: atvērtais BIM pret slēgto BIM. Eiropa un pārējā pasaule", 2024. gads. [Online]. Pieejams: <https://bigdataconstruction.com/lobbyist-wars-and-the-development-of-bim-part-2-open-bim-vs-closed-bim-revit-vs-archicad-and-europe-vs-the-rest-of-the-world/>.
  - [63] A. Boyko, "Lobbykriege um Daten im Bauwesen | Techno-Feudalismus und die Geschichte von BIMs," 2024. [Online]. Pieejams: [https://youtu.be/S-TNdUgfHxk?si=evM\\_v28KQbGOG0k&t=1360](https://youtu.be/S-TNdUgfHxk?si=evM_v28KQbGOG0k&t=1360).
  - [64] ADSK, "Baltā grāmata BIM", 2002. gads. [Online]. Pieejams:

- [https://web.archive.org/web/20060512180953/http://images.autodesk.com/apac\\_sapac\\_main/files/4525081\\_BIM\\_WP\\_Rev5.pdf#expand](https://web.archive.org/web/20060512180953/http://images.autodesk.com/apac_sapac_main/files/4525081_BIM_WP_Rev5.pdf#expand). [Adreses datums: 2025. gada 15. marts].
- [65] ADSK, "Integrated Design-Through-Manufacturing: Benefits and Rationale," [Online]. Pieejams: [https://web.archive.org/web/20010615093351/http://www3.adsk.com:80/adsk/files/734489\\_Benefits\\_of\\_MAI.pdf](https://web.archive.org/web/20010615093351/http://www3.adsk.com:80/adsk/files/734489_Benefits_of_MAI.pdf). [Piekļuves datums: 2025. gada 15. marts].
- [66] M. Shacklett, "Strukturēti un nestrukturēti dati: Galvenās atšķirības," 2024. [Online]. Pieejams: <https://www.datamation.com/big-data/structured-vs-unstructured-data/>.
- [67] K. Woolard, "Making sense of the growth of unstructured data", 2024. gads. [Online]. Pieejams: <https://automationhero.ai/blog/making-sense-of-the-rise-of-unstructured-data/>.
- [68] A. C. O. J. L. D. J. a. L. T. G. Michael P. Gallaher, "Neatbilstošas sadarbspējas izmaksu analīze", 2004. gads. [Online]. Pieejams: <https://nvlpubs.nist.gov/nistpubs/gcr/2004/nist.gcr.04-867.pdf>. [Piekļuves datums: 02 2025].
- [69] CrowdFlower, "Data Science Report 2016", 2016. gads. [Online]. Pieejams: [https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower\\_DataScienceReport\\_2016.pdf](https://visit.figure-eight.com/rs/416-ZBE-142/images/CrowdFlower_DataScienceReport_2016.pdf). [Piekļuves datums: 2025. gada 15. marts].
- [70] Analyticsindiamag, "6 datu zinātniekiem visvairāk laika patērējoši uzdevumi", 2019. gada 15. maijs. [Online]. Pieejams: <https://analyticsindiamag.com/ai-trends/6-tasks-data-scientists-spend-the-most-time-doing/>.
- [71] BizReport, "Ziņojums: datu zinātnieki lielāko daļu laika pavada tīrīšanai", 2015. gada 6. jūlijs. [Online]. Pieejams: <https://web.archive.org/web/20200824174530/http://www.bizreport.com/2015/07/report-data-scientists-spend-bulk-of-time-cleaning-up.html>. [Piekļuves datums: 2025. gada 5. marts].
- [72] S. Hawking, "Science AMA Series: Stephen Hawking AMA Answers!", 2015. gada 27. jūlijs. [Online]. Pieejams: [https://www.reddit.com/r/science/comments/3nyn5i/science\\_ama\\_series\\_stephen\\_hawking\\_a\\_ma\\_answers/](https://www.reddit.com/r/science/comments/3nyn5i/science_ama_series_stephen_hawking_a_ma_answers/). [Piekļuves datums: 2025. gada 15. marts].
- [73] B. Cyphers un K. Doctorow, "Privacy without Monopoly: Data Protection and Interoperability", 2024. [Online]. Pieejams: <https://www.eff.org/wp/interoperability-and-privacy>.
- [74] McKinsey Global Institute, "Open data: Unlocking innovation and performance with liquid information", 2013. gada 1. oktobris. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/open-data-unlocking-innovation-and-performance-with-liquid-information>. [Piekļuves datums: 2025. gada 15. marts].
- [75] A. Boyko, "Cīņa par atvērtiem datiem būvniecības nozarē. AUTOLISP, intelliCAD, openDWG, ODA

- un openCASCADE vēsture," 15 05 2024. [Online]. Pieejams: <https://boikoartem.medium.com/the-struggle-for-open-data-in-the-construction-industry-2b97200e6393>. [Adreses datums: 16 02 2025].
- [76] Vikipēdija, "Microsoft un atvērtā koda programmatūra", [Online]. Pieejams: [https://en.wikipedia.org/wiki/Microsoft\\_and\\_open\\_source](https://en.wikipedia.org/wiki/Microsoft_and_open_source). [Piekļuves datums: 2025. gada 15. marts].
- [77] TIME, "Atšķirība starp atvērtajiem un slēgtajiem mākslīgā intelekta modeļiem varētu samazināties. Lūk, kāpēc tas ir svarīgi," 2024. gada 5. novembris. [Online]. Pieejams: <https://time.com/7171962/open-closed-ai-models-epoch/>. [Piekļuves datums: 2025. gada 15. marts].
- [78] The Verge, "Vairāk nekā ceturtdaļu jaunā Google koda ģenerē mākslīgais intelekts," 29. oktobris, 2024. gada 29. oktobris. [Tiešsaistē]. Available: <https://www.theverge.com/2024/10/29/24282757/google-new-code-generated-ai-q3-2024>. [Piekļuves datums: 2025. gada 15. marts].
- [79] McKinsey Digital, "GPU izmantošanas biznesa pamatojums analītikas apstrādes paātrināšanai", 2020. gada 15. decembris. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/the-business-case-for-using-gpus-to-accelerate-analytics-apstrade>. [Piekļuve: 2025. gada 15. marts].
- [80] PwC, "PwC Open Source Monitor 2019", 2019. gads. [Online]. Pieejams: <https://www.pwc.de/de/digitale-transformation/open-source-monitor-research-report-2019.pdf>. [Skatīts 2025. gada 15. martā].
- [81] Travers Smith, "Atklāts noslēpums: atvērtā koda programmatūra", 2024. gads. [Online]. Pieejams: <https://www.traverssmith.com/knowledge/knowledge-container/the-open-secret-open-source-software/>. [Piekļuves datums: 2025. gada 15. marts].
- [82] Deloitte, "Datu pārsūtīšanas process uzņēmumu pārveidē", 2021. gads. [Online]. Pieejams: <https://www2.deloitte.com/content/dam/Deloitte/us/Documents/finance/us-the-data-transfer-process-in-corporate-transformations.pdf>. [Atsauces datums: 2025. gada 15. marts].
- [83] gov.uk, "Datu analītika un mākslīgais intelekts valdības projektu īstenošanā", 20.03.2024. [Online]. Pieejams: <https://www.gov.uk/government/publications/data-analytics-and-ai-in-government-project-delivery/data-analytics-and-ai-in-government-project-delivery>. [Piekļuves datums: 2025. gada 5. marts].
- [84] "Citāts "Izcelsme": Viss ir jāpadara pēc iespējas vienkāršāks, bet ne vienkāršāks," 2011. gada 13. maijs. [Online]. Pieejams: <https://quoteinvestigator.com/2011/05/13/einstein-simple/>. [Piekļuves datums: 2025. gada 15. marts].

- [85] "Transformators (dzīlās mācīšanās arhitektūra)," [Online]. Pieejams: [https://en.wikipedia.org/wiki/Transformer\\_\(deep\\_learning\\_architecture\)](https://en.wikipedia.org/wiki/Transformer_(deep_learning_architecture)). [Piekļuves datums: 2025. gada 15. marts].
- [86] "Python pakešu lejupielādes statistika", 2024. [Online]. Pieejams: <https://www.pypy.tech/projects/pandas>.
- [87] Intervija Bit, "Top 10 Python bibliotēkas", 2023. gads. [Tiešsaistē]. Pieejams: <https://www.interviewbit.com/blog/python-libraries/#:~:text=With%20more%20than%20137%2C000%20libraries,data%20manipulation%2C%20and%20many%20more>. [Piekļuves datums: 2025. gada 30. marts].
- [88] "NVIDIA un HP papildina datu zinātni un ģeneratīvo mākslīgo intelektu darbstacijās," 2025. gada 7. marts. [Online]. Pieejams: <https://nvidianews.nvidia.com/news/nvidia-hp-supercharge-data-science-generative-ai-workstations>. [Piekļuves datums: 2025. gada 15. marts].
- [89] P. Orac, "Kā dažu sekunžu laikā apstrādāt DataFrame ar miljoniem rindu", 2024. [Tiešsaistē]. Pieejams: <https://towardsdatascience.com/how-to-process-a-dataframe-with-millions-of-rows-in-seconds>.
- [90] Č. Uslu, "Kas ir Kaggle?", 2024. [Online]. Pieejams: <https://www.datacamp.com/blog/what-is-kaggle>.
- [91] "NVIDIA izpilddirektora Džensena Hjūanga (Jensen Huang) ievadvārdi izstādē COMPUTEX 2024", 2024. gada 2. jūnijs. [Online]. Pieejams: <https://www.youtube.com/live/pKXDVsWZmUU?si=Z3Rj1Las8wiPII2w>. [Skatīts 2025. gada 15. martā].
- [92] "Locekļi: dibinātāji un korporatīvie biedri", 2024. gads. [Online]. Pieejams: <https://www.opendesign.com/member-showcase>.
- [93] A. Boiko, "Pārmaiņu laikmets: IFC ir pagātne jeb kāpēc ADSK un citi CAD pārdevēji ir gatavi atteikties no IFC par USD, 14 galvenie fakti", 2024. gada 24. novembris. [Online]. Pieejams: <https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-are-willing-to-3f9a82ccd10a>. [Piekļuve: 2025. gada 23. februāris].
- [94] A. Boyko, "The post-BIM world. Pāreja uz datiem un procesiem, un vai būvniecības nozarei ir nepieciešama semantika, formāti un sadarbspēja," 20. dec. 2024. [Online]. Pieejams: <https://boikoartem.medium.com/the-post-bim-world-7e35b7271119>. [Uzrunas datums: 2025. gada 23. februāris].
- [95] N. I. o. Veselība, "NIH STRATĒĢISKAIS PLĀNS DATU ZINĀTNEI", 2016. gads. [Online]. Pieejams: [https://datascience.nih.gov/sites/default/files/NIH\\_Strategic\\_Plan\\_for\\_Data\\_Science\\_Final\\_508.pdf](https://datascience.nih.gov/sites/default/files/NIH_Strategic_Plan_for_Data_Science_Final_508.pdf). [Piekļuves datums: 2025. gada 23. februāris].

- [96] Harvard Business Review, "Slikti dati ASV izmaksā 3 triljonus ASV dolāru gadā", 2016. gada 22. septembris. [Online]. Pieejams: <https://hbr.org/2016/09/bad-data-costs-the-u-s-3-trillion-per-year>.
- [97] Delpha, "Datu kvalitātes ietekme", 2025. gada 1. janvāris. [Online]. Pieejams: <https://delpha.io/impacts-of-data-quality/>.
- [98] W. B. D. Guide, "Design for Maintainability: The Importance of Operations and Maintenance Considerations During the Design Phase of Construction Projects" ("Projektēšana uzturēšanai: ekspluatācijas un uzturēšanas apsvērumu nozīme būvniecības projektu projektēšanas posmā"), [Online]. Pieejams: <https://www.wbdg.org/resources/design-for-maintainability>. [Piekļuves datums: 2025. gada 15. marts].
- [99] O. o. D. C. P. a. Oversight, "Korozijas novēršanas un kontroles plānošanas rokasgrāmata militārajām sistēmām un iekārtām", 2014. gada aprīlis. [Online]. Pieejams: <https://www.dau.edu/sites/default/files/Migrated/CopDocuments/CPC%20Planning%20Guide%204%20Feb%202014.pdf>. [Piekļuves datums: 2025. gada 15. marts].
- [100] Gartner, "Datu kvalitāte: labākā prakse precīzai informācijai", 2025. gada 1. janvāris. [Online]. Pieejams: <https://www.gartner.com/en/data-analytics/topics/data-quality>.
- [101] "Vainaga trūkuma dēļ," [Online]. Pieejams: [https://en.wikipedia.org/wiki/For\\_Want\\_of\\_a\\_Nail](https://en.wikipedia.org/wiki/For_Want_of_a_Nail). [Piekļuves datums: 2025. gada 15. marts].
- [102] McKinsey Global Institute, "Open data: Unlocking innovation and performance with liquid information", 2013. gada oktobris. [Online]. Pieejams: [https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/open%20data%20unlocking%20innovation%20and%20performance%20with%20liquid%20information/mgi\\_open\\_data\\_fullreport\\_oct2013.pdf](https://www.mckinsey.com/~media/mckinsey/business%20functions/mckinsey%20digital/our%20insights/open%20data%20unlocking%20innovation%20and%20performance%20with%20liquid%20information/mgi_open_data_fullreport_oct2013.pdf). [Piekļuves datums: 2025. gada 15. marts].
- [103] EY, "Ceļš uz oglekļa neutralitāti", 2023. gada 10. marts. [Online]. Pieejams: [https://www.ey.com/ru\\_kz/services/consulting/the-path-to-carbon-neutrality](https://www.ey.com/ru_kz/services/consulting/the-path-to-carbon-neutrality). [Piekļuves datums: 2025. gada 15. marts].
- [104] PWC, "ESG izpratne", 2024. gada 1. jūlijs. [Online]. Pieejams: <https://www.pwc.com/kz/en/assets/esg-awareness/kz-esg-awareness-rus.pdf>. [Piekļuves datums: 2025. gada 15. marts].
- [105] G. Hamonds, "Embodied Carbon - The Inventory of Carbon and Energy (ICE)", 2024. gads. [Online]. Pieejams: <https://greenbuildingencyclopaedia.uk/wp-content/uploads/2014/07/Full-BSRIA-ICE-guide.pdf>.
- [106] "CO2\_rēķini, kuros ietverts ogleklis", 2024. gads. [Online]. Pieejams:

- [https://github.com/datadrivenconstruction/CO2\\_calculating-the-embodied-carbon](https://github.com/datadrivenconstruction/CO2_calculating-the-embodied-carbon).
- [107] McKinsey, "Imagining Construction's Digital Future", 2016. gada 24. jūnijs. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/operations/our-insights/imagining-constructions-digital-future>. [Piekļuves datums: 2025. gada 25. februāris].
- [108] Bund der Steuerzahler Deutschland e.V., "Das Schwarzbuch", 2024. gada 10. oktobris. [Online]. Pieejams: <https://steuerzahler.de/aktuelles/detail/das-schwarzbuch-202425/>. [Adreses datums: 2025. gada 15. marts].
- [109] SAS, "Datu ezers un datu glabātuve - izprotiet atšķirību," [Online]. Pieejams: [https://www.sas.com/en\\_is/insights/articles/data-management/data-lake-and-data-warehouse-know-the-difference.html](https://www.sas.com/en_is/insights/articles/data-management/data-lake-and-data-warehouse-know-the-difference.html). [Atsauces datums: 2025. gada 15. marts].
- [110] ADSK, "Building Information Modeling", 2002. gads. [Online]. Pieejams: [https://www.laiserin.com/features/bim/autodesk\\_bim.pdf](https://www.laiserin.com/features/bim/autodesk_bim.pdf). [Adreses datums: 2025. gada 15. marts].
- [111] A. Boiko, "BIM vēstures karte", 2024. [Online]. Pieejams: <https://bigdataconstruction.com/history-of-bim/>.
- [112] A. S. Borkovskis, "BIM definīcijas pēc organizācijām un standartiem", 2023. gada 27. decembris. [Online]. Pieejams: <https://encyclopedia.pub/entry/53149>. [Adreses datums: 2025. gada 5. marts].
- [113] CAD pārdevējs, "OPEN BIM programma", 2012. gads. [Online]. Pieejams: [https://web.archive.org/web/20140611075601/http://www.graphisoft.com/archicad/open\\_bim/](https://web.archive.org/web/20140611075601/http://www.graphisoft.com/archicad/open_bim/). [Atsauces datums: 2025. gada 30. marts].
- [114] Vikipēdija, "Nozares pamatklases," [Online]. Pieejams: [https://en.wikipedia.org/wiki/Industry\\_Foundation\\_Classes](https://en.wikipedia.org/wiki/Industry_Foundation_Classes). [Piekļuves datums: 2025. gada 15. marts].
- [115] Vikipēdija, "IGES," [Online]. Pieejams: <https://en.wikipedia.org/wiki/IGES>. [Atsauces datums: 2025. gada 30. marts].
- [116] A. Boiko, "CAD (BIM) vēsture", 2021. gada 15. decembris. [Online]. Pieejams: [https://miro.com/app/board/o9J\\_laML2cs=/](https://miro.com/app/board/o9J_laML2cs=/). [Adreses datums: 2025. gada 24. februāris].
- [117] T. K. K. A. O. F. B. C. E. L. H. H. H. E. L. P. N. S. H. H. T. J. v. L. H. H. G. D. H. T. K. C. L. A. W. J. S. Francesca Noardo, "IFC programmatūras atbalsta standartpētījums: GeoBIM etalons 2019 - I daļa", 8. janvāris 2021. gadā. [Online]. Pieejams: <https://arxiv.org/pdf/2007.10951>. [Piekļuves datums: 2025. gada 5. marts].

- [118] И. Rogačevs, "Parunāsim par BIM: Maksims Nečiporenko | Renga | IFC | Domestic BIM", 2021. gada 13. aprīlis. [Online]. Pieejams: <https://www.youtube.com/watch?t=3000&v=VO3Y9uuzF9M&feature=youtu.be>. [Uzrunas datums: 2025. gada 5. marts].
- [119] D. Ares, "RETS nekustamajā īpašumā: kāpēc tas ir izšķirošs efektivitātei un izaugsmei", 2024. gada 17. decembris. [Online]. Pieejams: <https://www.realpha.com/blog/rets-importance-in-real-estate-explained>. [Piekļuves datums: 2025. gada 5. marts].
- [120] "Flex token izmaksas", 2024. gads. [Online]. Pieejams: <https://www.adsk.com/buying/flex?term=1-YEAR&tab=flex>.
- [121] A. Boyko, "Aizmirstiet BIM un demokratizējet piekļuvi datiem (17. Kolloquium Investor - Hochschule - Bauindustrie)," 2024. [Online]. Pieejams: <https://www.bim.bayern.de/wp-content/uploads/2023/06/Kolloquium-17-TUM-Bauprozessmanagment-und-Bauindustrie.pdf>.
- [122] Д. Hill, D. Foldesi, S. Ferrer, M. Friedman, E. Loh un F. Plaschke, "Solving the construction industry productivity puzzle", 2015. gads. [Online]. Pieejams: <https://www.bcg.com/publications/2015/engineered-products-project-business-solving-construction-industrys-productivity-puzzle>.
- [123] "SCOPE - Projektdatenumgebung und Modellierung multifunktionaler Bauprodukte mit Fokus auf die Gebäudehülle," 2018. gada 1. janvāris. [Online]. Pieejams: <https://www.ise.fraunhofer.de/de/forschungsprojekte/scope.html>. [Piekļuves datums: 2025. gada 2. marts].
- [124] Apple.com, "Pixar, Adobe, Apple un NVIDIA veido aliansi OpenUSD, lai veicinātu atvērtus 3D satura standartus", 2023. gada 1. augusts. [Tiešsaistē]. Pieejams: <https://www.apple.com/newsroom/2023/08/pixar-adobe-apple-adsk-and-nvidia-form-alliance-for-openusd/>. [Piekļuve: 2025. gada 2. marts].
- [125] AECmag, "ADSK granulāro datu stratēģija", 25 Jul 2024. [Online]. Pieejams: <https://aecmag.com/technology/autodesk-granular-data-strategy/>. [Piekļuves datums: 2025. gada 15. marts].
- [126] A. Boiko, "Pārmaiņu laikmets: IFC ir pagātne jeb kāpēc ADSK un citi CAD ražotāji ir gatavi atteikties no IFC par USD 14 galvenie fakti", 24 11 2024. [Online]. Pieejams: <https://boikoartem.medium.com/the-age-of-change-ifc-is-a-thing-of-the-past-or-why-adsk-and-other-cad-vendors-are-willing-to-3f9a82cccd10a>. [Adreses datums: 2025. gada 23. februāls].
- [127] A. Boyko, "ENG BIM klasteris 2024 | Cīņa par datiem un LLM un ChatGPT pielietojumu būvniecībā", 2024. gada 7. augusts. [Online]. Pieejams: ENG BIM Cluster 2024 | The Battle for Data and Application of LLM and ChatGPT in the Construction. [Piekļuve: 2025. gada 15. marts].

- [128] "Džefrijs Zeldmens iepazīstina", 2008. gada 6. maijs. [Online]. Pieejams: <https://zeldman.com/2008/05/06/content-precedes-design/>. [Piekļuves datums: 2025. gada 15. marts].
- [129] A. Boyko, "DWG analīze ar ChatGPT | DataDrivenConstruction", 2024. gada 5. marts. [Tiešsaistē]. Pieejams: <https://www.kaggle.com/code/artemboiko/dwg-analyse-with-chatgpt-datadrivenconstruction>. [Atsauces datums: 2025. gada 15. marts].
- [130] McKinsey , "McKinsey ceļvedis, kā pārspēt konkurentus digitalizācijas un mākslīgā intelekta laikmetā", 2023. gads. [Online]. Pieejams: <https://www.mckinsey.com/featured-insights/mckinsey-on-books/rewired>. [Adreses datums: 2025. gada 30. marts].
- [131] Forbes, "Datu stāstu stāstīšana: būtiska datu zinātnes prasme, kas nepieciešama ikvienam", 2016. gada 31. marts. [Online]. Pieejams: <https://www.forbes.com/sites/brentdykes/2016/03/31/data-storytelling-the-essential-data-science-skill-everyone-needs/>. [Piekļuves datums: 2025. gada 15. marts].
- [132] J. Bertin, "Grafika un grafiskā informācijas apstrāde", 2011. gada 8. septembris. [Online]. Pieejams: [https://books.google.de/books/about/Graphics\\_and\\_Graphic\\_Information\\_Process.html?id=csqX\\_xnm4tcC&redir\\_esc=y](https://books.google.de/books/about/Graphics_and_Graphic_Information_Process.html?id=csqX_xnm4tcC&redir_esc=y). [Piekļuves datums: 2025. gada 15. marts].
- [133] CauseWeb, "Wells/Wilks on Statistical Thinking," [Online]. Pieejams: <https://www.causeweb.org/cause/resources/library/r1266>. [Piekļuves datums: 2025. gada 15. marts].
- [134] Ministrymagazine, "Kā zinātne atklāja radīšanu", 1986. gada janvāris. [Online]. Pieejams: <https://www.ministrymagazine.org/archive/1986/01/how-science-discovered-creation>. [Atsauces datums: 2025. gada 15. marts].
- [135] BCG, "Data-Driven Transformation: Accelerate at Scale Now", 2017. gada 23. maijs. [Online]. Pieejams: <https://www.bcg.com/publications/2017/digital-transformation-transformation-data-driven-transformation>. [Skatīts 2024. gada 15. maijā].
- [136] "Kā izveidot datu arhitektūru, lai veicinātu inovācijas šodien un rīt," 2020. gada 3. jūnījs. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/how-to-build-a-data-architecture-to-drive-innovation-today-and-tomorrow>. [Uzrunas datums: 2025. gada 15. marts].
- [137] Oksforda, "Woodrow Wilson 1856-1924," [Online]. Available: <https://www.oxfordreference.com/display/10.1093/acref/9780191866692.001.0001/q-oro-ed6-00011630>. [Uzrunas datums: 2025. gada 15. marts].
- [138] "Konvertori", 2024. [Online]. Pieejams: <https://datadrivenconstruction.io/index.php/convertors/>.

- [139] PWC, "Sizing the prize What's the real value of AI for your business and how you can kapitalizēt?", 2017. gada 1. janvāris. [Online]. Pieejams: <https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf>. [Piekļuves datums: 2025. gada 18. februāris].
- [140] "Cauruļvads būvniecībā", 2024. gads. [Online]. Pieejams: <https://datadrivenconstruction.io/index.php/pipeline-in-construction/>.
- [141] Wikipedia, "Apache NiFi", 2025. gada 1. janvāris. [Online]. Pieejams: [https://de.wikipedia.org/wiki/Apache\\_NiFi](https://de.wikipedia.org/wiki/Apache_NiFi). [Piekļuves datums: 2025. gada 5. marts].
- [142] n8n, "Gmail AI automātiskais atbildētājs: atbildes projektu izveide uz ienākošajiem e-pasta ziņojumiem", 2024. gada 1. maijs. [Tiešsaistē]. Pieejams: <https://n8n.io/workflows/2271-gmail-ai-auto-responder-create-draft-replies-to-incoming-emails/>. [Skatīts 2025. gada 15. martā].
- [143] n8n, "Real Estate Daily Deals Automation with Zillow API, Google Sheets and Gmail", 2025. gada 1. marts. [Tiešsaistē]. Pieejams: <https://n8n.io/workflows/3030-real-estate-daily-deals-automation-with-zillow-api-google-sheets-and-gmail/>. [Piekļuves datums: 2025. gada 15. marts].
- [144] B. T. O'Nīls, "Analītisko, mākslīgā intelekta un lielo datu projektu neveiksmju rādītāji = 85 % - yikes!", 2025. gada 1. janvāris. [Online]. Pieejams: <https://designingforanalytics.com/resources/failure-rates-for-analytics-bi-iot-and-big-data-projects-85-yikes/>.
- [145] J. Neyman, On the Two Different Aspects of the Representative Method: The Method of Stratified Sampling and the Method of Purposive Selection, Oxford University Press, 1934. gadā.
- [146] T. J. S. a. J. S. Jesse Perla, "A Problem that Stumped Milton Friedman," Quantitative Economics with Julia, 2025. gada 1. janvāris. [Online]. Pieejams: [https://julia.quantecon.org/dynamic\\_programming/wald\\_friedman.html](https://julia.quantecon.org/dynamic_programming/wald_friedman.html). [Piekļuves datums: 2024. gada 1. maijs].
- [147] T. Landsall-Welfair, "Prognozējot nācijas pašreizējo noskaņojumu", Significance, 2012. gads.
- [148] A. Boiko, "Sanfrancisko. Būvniecības nozare 1980-2019," 2024. [Online]. Pieejams: <https://www.kaggle.com/search?q=San+Francisco.+Building+sector+1980-2019>.
- [149] A. Boyko, "Kaggle: RVT IFC iesniedz 5000 projektu", 2024. [Online]. Pieejams: <https://www.kaggle.com/datasets/artemboiko/rvtifc-projects>.
- [150] CFMA, "Gatavojoties nākotnei ar savienoto būvniecību". [Online]. Pieejams: <https://cfma.org/articles/preparing-for-the-future-with-connected-construction>. [Skatīts 2025].

- gada 15. martā].
- [151] Cisco, "Cisco aptauja atklāj, ka gandrīz trīs ceturtdaļas IoT projektu cieš neveiksmi", 2017. gada 22. maijs. [Online]. Pieejams: <https://newsroom.cisco.com/c/r/newsroom/en/us/a/y2017/m05/cisco-survey-reveals-close-to-three-fourths-of-iot-projects-are-failing.html>.
- [152] "Augu fosiliju saglabāšanai nepieciešamie apstākļi", 2024. gads. [Online]. Pieejams: <https://ucmp.berkeley.edu/IB181/VPL/Pres/PresTitle.html>.
- [153] "BlackRock's Fink par obligācijām, uzņēmumu apvienošanu un pārņemšanu, ASV recesiju un vēlēšanām: Pilna intervija", 2023. gads. [Online]. Pieejams: <https://www.bloomberg.com/news/videos/2023-09-29/blackrock-s-fink-on-m-a-recession-election-full-intv-video>.
- [154] cio, "12 slavenas mākslīgā intelekta katastrofas", 2024. gada 02. oktobris. [Online]. Pieejams: <https://www.cio.com/article/190888/5-famous-analytics-and-ai-disasters.html>. [Piekļuves datums: 2025. gada 15. marts].
- [155] G. Kasparovs, Deep Thinking, PublicAffairs, 2017.
- [156] Vikipēdija, "Kaggle", 2025. gada 1. janvāris. [Tiešsaistē]. Pieejams: <https://en.wikipedia.org/wiki/Kaggle>. [Piekļuves datums: 2025. gada 15. marts].
- [157] Kaggle, "Titāniks - mašīnmācīšanās no katastrofas", 2025. gada 1. janvāris. [Tiešsaistē]. Pieejams: <https://www.kaggle.com/competitions/titanic/overview>. [Skatīts 2025. gada 10. martā].
- [158] Ш. Johri, "Creating ChatGPT: No datiem līdz dialogam", 2024. [Online]. Pieejams: <https://sitn.hms.harvard.edu/flash/2023/the-making-of-chatgpt-from-data-to-dialogue/>.
- [159] П. Domingos, "Dažas noderīgas lietas, kas jāzina par mašīnmācīšanos", 2024. gads. [Online]. Pieejams: <https://homes.cs.washington.edu/~pedrod/papers/cacm12.pdf>.
- [160] J. Saramago, "Citējams citāts", [Online]. Pieejams: <https://www.goodreads.com/quotes/215253-chaos-is-merely-order-waiting-to-be-deciphered>. [Piekļuves datums: 2025. gada 17. marts].
- [161] NVIDIA, "Uzlabojet mācību datus ar jaunajiem NVIDIA NeMo Curator klasifikatoru modeļiem", 2024. gada 19. decembris. [Tiešsaistē]. Pieejams: <https://developer.nvidia.com/blog/enhance-your-training-data-with-new-nvidia-nemo-curator-classifier-models/>. [Skatīts 2025. gada 25. martā].
- [162] "NVIDIA paziņo par Cosmos World Foundation modeļu un fizikālo AI datu rīku būtisku versiju,"

2025. gada 18. marts. [Online]. Pieejams: <https://nvidianews.nvidia.com/news/nvidia-announces-major-release-of-cosmos-world-foundation-models-and-physical-ai-data-tools>. [Piekļuves datums: 2025. gada 25. marts].
- [163] NVIDIA, "NVIDIA Isaac Sim," [Internetā]. Pieejams: <https://developer.nvidia.com/isaac/sim>. [Skatīts 2025. gada 25. martā].
- [164] M. Quarterly, "Why digital strategies fail," 25 Jan. 2018. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-digital-strategies-fail>. [Piekļuves datums: 2025. gada 15. marts].
- [165] M. J. Perry, "Mani mīlākie Miltona Frīdmēna citāti", 2006. gada 17. novembris. [Online]. Pieejams: <https://www.aei.org/carpe-diem/my-favorite-milton-friedman-quotes/>. [Piekļuves datums: 2025. gada 1. marts].
- [166] J. A. Wheeler, "Informācija, fizika, kvanti: saikņu meklējumi", 1990. gads.
- [169] A. Boyko, "Lobiju kari un BIM attīstība. 5. daļa: BlackRock ir visu tehnoloģiju meistars. Kā korporācijas kontrolē atvērtā pirmkoda kodus," 2024. gads. [Online]. Pieejams: <https://boikoartem.medium.com/lobbyist-wars-and-the-development-of-bim-d72ad0111a7d>.
- [170] T. Krijnen un J. Beetz, "A SPARQL query engine for binary-formatted IFC building models," *Advanced Engineering Informatics*, 2024.
- [171] "Uzņēmumu skaits būvniecības nozarē Apvienotajā Karalistē 2021. gadā pēc uzņēmumu lieluma", 2024. gads. [Online]. Pieejams: <https://www.statista.com/statistics/677151/uk-construction-businesses-by-size/>.
- [172] "5000 IFC&RVT projekti," 2024. gads. [Online]. Pieejams: <https://www.kaggle.com/code/artemboiko/5000-projects-ifc-rvt-datadrivenconstruction-io>.
- [173] M. Popova, "It from Bit: Pioneering Physicist John Archibald Wheeler on Information, the Nature of Reality, and Why We Live in a Participatory Universe," 2008. gads. [Online]. Pieejams: <https://www.themarginalian.org/2016/09/02/it-from-bit-wheeler/>. [Atsauces datums: 2025. gada februāris].
- [174] *Lobiju kari par datiem būvniecībā / Tehnofeodālisms un BIM slēptās pagātnes vēsture*. [Filma]. Vācija: Artyom Boiko, 2023.
- [175] A. Boyko, "CHATGPT WITH REVIT AND IFC | Automātiska dokumentu un datu izgūšana no projektiem", 2023. gada 16. novembris. [Tiešsaistē]. Pieejams: [https://www.youtube.com/watch?v=ASXolti\\_YPs&t](https://www.youtube.com/watch?v=ASXolti_YPs&t). [Piekļuves datums: 2025. gada 2. marts].
- [176] M. & Uzņēmums, "Trīs jauni mandāti, lai pilnībā izmantotu digitālās transformācijas vērtību",

2022. gada 22. janvāris. [Online]. Pieejams: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/three-new-mandates-for-capturing-a-digital-transformations-full-value>. [Piekļuves datums: 2025. gada 15. februāris].
- [177] KPMG, "Būvniecība digitālajā pasaule", 2021. gada 1. maijs. [Online]. Pieejams: <https://assets.kpmg.com/content/dam/kpmg/ca/pdf/2021/05/construction-in-the-digital-age-report-en.pdf>. [Piekļuves datums: 2024. gada 5. aprīlis].
- [178] LLP, KPMG, "Cue Construction 4.0: Make-or-Break Time", 2023. gada 17. marts. [Online]. Pieejams: <https://kpmg.com/ca/en/home/insights/2023/05/cue-construction-make-or-break-time.html>. [Piekļuves datums: 2025. gada 15. februāris].
- [179] O. Bizness, "Satya Nadella atklāj, kā mākslīgā intelekta aģenti sagraus SaaS modeļus", 2025. gada 10. janvāris. [Online]. Pieejams: <https://www.outlookbusiness.com/artificial-intelligence/microsoft-ceo-satya-nadella-reveals-how-ai-agents-will-disrupt-saas-models>. [Piekļuves datums: 2025. gada 15. marts].
- [180] Forbes, "Lielo datu tīrišana: vislaikītilpīgākais un vismazāk patīkamas datu zinātnes uzdevums," Forbes, "Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says", 2016. gada 23. marts. [Online]. Pieejams: <https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/>. [Piekļuves datums: 2025. gada 15. marts].
- [181] Apvienotās Karalistes Ārlietu, Sadraudzības un attīstības ministrija, "Digitālās attīstības stratēģija 2024-2030", 18.03.2024. [Online]. Pieejams: <https://www.gov.uk/government/publications/digital-development-strategy-2024-to-2030/digital-development-strategy-2024-to-2030>. [Piekļuves datums: 2025. gada 15. marts].
- [182] "Vīzija un stratēģija ēku projektēšanas nozarē", 2003. gada 7. novembris. [Online]. Pieejams: <https://web.archive.org/web/20030711125527/http://usa.adsk.com/adsk/servlet/item?id=2255342&siteID=123112>. [Uzrunas datums: 2025. gada 5. marts].
- [183] M. Bočarovs, "Informācijas modelēšana", 2025. gada marts. [Online]. Pieejams: <https://www.litres.ru/book/mihail-evgenievich-bocharov/informacionnoe-modelirovanie-v-rossii-71780080/chitat-onlayn/?page=5>. [Uzrunas datums: 2025. gada 15. marts].
- [184] "Integrēta projektēšana caur ražošanu: ieguvumi un pamatojums", 2000. gads. [Online]. Pieejams: [https://web.archive.org/web/20010615093351/http://www3.autodesk.com:80/adsk/files/734489\\_Benefits\\_of\\_MAI.pdf](https://web.archive.org/web/20010615093351/http://www3.autodesk.com:80/adsk/files/734489_Benefits_of_MAI.pdf). [Atsauces datums: 2025. gada 25. marts].
- [185] CAD Vendor, "Atvērtā BIM programma ir mārketinga kampanja," 2012. gada 12. marts. [Online]. Pieejams: <https://web.archive.org/web/20120827193840/http://www.graphisoft.com/openbim/>. [Adreses

datums: 2025. gada 30. marts].

# SUBJEKTU INDEKSS

3D, 8, 14, 71, 73, 84, 191, 210, 215, 232, 234, 263, 276, 277, 287, 298, 299, 302, 303, 306, 307, 337, 375, 393, 394, 448, 468, 480

4D, 84, 172, 196, 199, 210, 229, 234, 237, 287  
4IR, 43

5D, 84, 172, 196, 210, 237, 287, 492

6D, 172, 196, 229, 232, 233, 234, 235

7D, 232, 233, 234, 287

8D, 172, 196, 229, 232, 233, 234, 287

## A

AI, 3, 50, 52, 100, 102, 103, 106, 107, 116, 395, 457, 459, 461, 477

AIA, 289

AIM, 3, 289

AMS, 14, 84, 151, 153

Apache Airflow, 311, 361, 362, 363, 364, 366, 367, 369, 371, 399, 494

Apache NiFi, 116, 311, 361, 367, 368, 371, 399, 482, 494

Apache ORC, 62, 330, 378

Apache Parquet, 62, 67, 380, 381, 494

API, 54, 94, 95, 97, 109, 137, 138, 154, 168, 218, 219, 256, 257, 260, 271, 273, 294, 295, 296, 297, 300, 308, 326, 329, 342, 366, 369, 370, 488

## B

BDS, 258, 259

Lielie dati, 9, 67, 245, 383

BIM, 2, 3, 4, 6, 3, 14, 17, 24, 56, 58, 60, 63, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 84, 92, 137, 138, 139, 140, 141, 142, 144, 150, 154, 156, 166, 172, 183, 184, 186, 187, 190, 195, 196, 208, 210, 211, 213, 216, 217, 218, 219, 220, 221, 222, 227, 228, 237, 238, 239, 242, 243, 246, 250, 251, 252, 255, 256, 257, 258, 259, 260, 261, 262, 263, 266, 271, 272, 273, 275, 276, 277, 278, 279, 280, 282, 285, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 308, 309, 324, 328, 337, 351, 356, 358, 361, 367, 372, 376, 383, 389, 402, 403, 405, 406, 413, 416, 431, 448, 462, 466, 475, 476, 480, 492, 494, 497

475, 476, 492, 494, 497

BlackBox, 240, 242, 243.

BMS, 8

Bokeh, 320, 337

BOM, 76, 77, 79, 257, 263

Ierobežojošais lodziņš, 234, 373, 392, 393, 394, 414, 415, 480

BREP, 142, 234, 263, 264, 276, 283, 284

## C

KAD, 6, 14, 18, 24, 56, 57, 58, 63, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 84, 85, 95, 95, 97, 111, 126, 137, 138, 139, 140, 141, 142, 144, 146, 147, 152, 153, 155, 156, 166, 172, 175, 183, 184, 186, 187, 190, 195, 196, 206, 208, 210, 211, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 224, 227, 228, 232, 234, 237, 238, 239, 242, 243, 251, 252, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 303, 308, 309, 324, 328, 337, 344, 351, 356, 358, 361, 367, 372, 376, 383, 389, 402, 403, 405, 406, 413, 416, 431, 448, 462, 466, 475, 476, 480, 492, 494, 497

CAE, 16, 283

CAFM, 14, 24, 62, 84, 151, 153, 172, 233, 278, 326, 356, 387, 458, 473

CAM, 16, 78

CAPEX, 14, 82

CDE, 84, 175, 388, 389, 390

ChatGPT, 103, 104, 109, 110, 124, 129, 221, 303, 307, 343, 349, 421

Klods, 103, 104, 109, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 429, 443.

CO<sub>2</sub>, 72

CO<sub>2</sub>, 229, 234, 235, 236, 237, 238, 239, 243

COBie, 156, 289, 292.

EP, 56, 168, 169, 169, 170, 477

otrais pilots, 114, 116, 459

CPIXML, 143, 272, 273, 276, 277, 279, 285, 296

CPM, 14, 17, 62, 166, 175, 233, 326, 473

CQMS, 14, 84, 177, 178, 462

CRM, 109, 369, 459.

CRUD, 51, 459

CSG, 263

CSV, 61, 62, 88, 89, 120, 123, 128, 129, 130, 131, 135, 168, 268, 272, 280, 329, 333, 344, 354, 354, 356, 373, 377, 378, 379, 380, 401, 407, 414, 474, 480

**D**

DAE, 276, 277, 278, 280, 281, 284, 285, 414  
 DAG, 362, 363, 365, 366  
 Dash, 320, 336, 337.  
 Datu pārvaldība, 373, 395, 396, 398, 400, 401  
 Datu ezers, 214, 373, 376, 384, 385, 386, 387, 388, 389, 390, 400  
 Data Lakehouse, 67, 373, 386, 387, 387, 388  
 Datu minimālisms, 373, 395, 396, 400, 401  
 Datu purvs, 373, 395, 397, 401  
 Datu glabātuve, 382, 383, 400  
 Dati kā pakalpojums, 487  
 uz datiem balstīta, 50, 170, 460, 461, 484, 486  
 DataFrame, 67, 117, 121, 122, 123, 125, 129, 130, 131, 133, 134, 135, 137, 219, 220, 224, 237, 308, 328, 329, 330, 332, 333, 344, 345, 347, 348, 354, 365, 377, 381, 409, 414, 415, 475  
 DataOps, 170, 373, 398, 399, 400, 401, 475  
 DeepSeek, 103, 104, 107, 109, 110, 124, 129, 133, 166, 219, 221, 300, 329, 343, 349, 349, 381, 421, 429, 443, 494  
 DGN, 8, 140, 186, 227, 357  
 DWG, 8, 70, 71, 73, 97, 140, 186, 211, 227, 272, 287, 302, 303, 304, 307, 357, 376, 497  
 DWH, 67, 373, 376, 382, 383, 384, 386, 387, 388, 389, 390  
 DXF, 8, 73, 277

**E**

ECM, 58, 175  
 ECS, 142  
 EIR, 289  
 eLOD, 289  
 ELT, 384, 385  
 EPM, 14, 166, 198  
 ERP, 2, 11, 12, 14, 17, 18, 24, 25, 58, 62, 109, 153, 166, 172, 175, 196, 198, 210, 232, 239, 240, 241, 242, 243, 244, 245, 246, 247, 249, 272, 277, 278, 279, 282, 326, 351, 356, 361, 369, 387, 388, 389, 390, 458, 462, 473, 475, 487, 492  
 ESG, 196, 235, 236, 238  
 ETL, 6, 1.1-8, 32, 49, 81, 113, 116, 119, 128, 188, 193, 219, 291, 311, 312, 317, 323, 324, 325, 326, 327, 328, 330, 331, 333, 338, 339, 340, 343, 344, 348, 349, 350, 351, 353, 354, 356, 361, 362, 363, 364, 365, 367, 371, 372, 381, 382, 383, 384, 385, 399, 409, 430, 475, 481, 494  
 Excel, 57, 61, 62, 65, 66, 85, 88, 111, 120, 123, 123, 125, 132, 154, 167, 187, 210, 223, 224, 226, 227, 228, 278, 291, 329, 333, 342, 351, 365, 376, 378, 459, 475, 497.  
 Izraksts, 81, 128, 134, 193, 311, 323, 324, 326, 328, 330, 345, 361, 362, 363, 365, 383, 384, 450, 475.

**F**

Spalvu, 62, 123, 330, 378  
 FPDF, 339, 340, 341, 342, 343  
**G**  
 GDPR, 109  
 GIS, 58  
 GLTF, 143, 278  
 Google lapas, 368, 370  
 Grok, 103, 104, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 429, 443

**H**

HDF5, 62, 67, 123, 329, 330, 378, 379, 380  
 HiPPO, 29, 37, 95, 424, 477, 484, 490  
 HTML, 123, 340, 365, 370

**I**

IDS, 289, 290, 291.  
 IFC, 8, 73, 138, 142, 186, 227, 261, 262, 263, 264, 265, 266, 267, 268, 272, 273, 273, 276, 277, 278, 279, 280, 284, 286, 292, 296, 302, 329, 357, 414, 415, 417, 497  
 IGES, 262, 263, 276.  
 iLOD, 289  
 IoT, 10, 18, 67, 271, 367, 369, 405, 413, 417, 417, 418, 419, 455, 460, 465, 482, 484.  
 ISO 19650, 388

**J**

JavaScript, 320, 369, 378  
 JSON, 88, 89, 90, 92, 123, 128, 142, 269, 272, 280, 329, 330, 333, 378, 474, 480  
 Jupyter Notebook, 114, 115, 116, 130, 187, 224, 330, 346, 417, 425

**K**

Kaggle, 115, 121, 130, 187, 224, 303, 307, 330, 346, 408, 415, 417, 425, 426, 430, 431, 433  
 k-NN, 392, 393, 442, 445, 446, 447, 448  
 KPI, 245, 311, 317, 318, 319, 320, 321, 324, 353, 372, 389, 478, 479

**L**

LEED, 235, 236, 238.  
 LlaMa, 103, 104, 120, 124, 129, 133, 166, 219, 300, 329,

343, 349, 381, 421, 429, 443, 494

LLM, 3, 4, 24, 29, 50, 51, 52, 55, 56, 92, 95, 99, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 118, 120, 124, 125, 126, 129, 131, 133, 134, 135, 166, 166, 187, 219, 220, 220, 221, 222, 222, 224, 225, 226, 231, 237, 238, 251, 294, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 328, 329, 332, 333, 334, 335, 336, 338, 339, 340, 343, 344, 345, 346, 347, 348, 349, 354, 370, 372, 376, 381, 382, 392, 394, 399, 413, 425, 429, 430, 432, 433, 435, 443, 444, 457, 460, 461, 476, 488, 494  
Slodze, 81, 128, 134, 193, 311, 323, 325, 326, 333, 334, 338, 339, 342, 343, 361, 362, 363, 365, 367, 383, 384, 450, 475.  
LOD, 287, 289  
LOI, 287  
LOMD, 287  
Low-Code, 368, 369

## M

Matplotlib, 123, 306, 320, 333, 335, 342, 372, 409, 411, 494  
MCAD, 77, 78, 257, 284  
MEP, 14, 175, 492  
SIETS, 234, 283, 284, 285, 296, 462  
Microsoft SQL, 65  
Mistral, 103, 104, 107, 110, 129, 133, 166, 219, 221, 300, 329, 343, 349, 381, 421, 429, 443, 494  
MRP, 11, 12  
MS Project, 70  
MySQL, 63, 64, 65, 291, 329, 330

## N

n8n, 116, 311, 361, 368, 369, 370, 371, 372  
NLP, 69, 394  
No-Code, 368, 369  
NURBS, 142, 282, 283, 283, 284, 285  
NWC, 8, 276

## O

OBJ, 143, 273, 276, 277, 278, 280, 281, 284, 285  
OCCT, 273  
OCR, 69, 128, 132, 134  
OmniClass, 154, 155, 156.  
atvērtais BIM, 142, 216, 219, 256, 261, 278, 291  
Atvērtā pirmkoda, 45, 55, 97, 98, 107, 108, 273, 275, 494  
PŪCE, 267, 268, 269.

## P

Pandas, 56, 67, 103, 117, 118, 119, 120, 121, 122, 123, 125, 130, 134, 138, 186, 187, 220, 221, 225, 227, 269, 300, 303, 328, 329, 330, 377, 378, 380, 381, 403, 408, 409, 410, 411, 412, 414, 475, 479, 494  
Parkets, 67, 123, 329, 330, 373, 378, 379, 380, 381, 382, 386, 401, 403, 414, 474, 480  
PDF, 69, 70, 71, 85, 111, 126, 127, 128, 129, 130, 131, 132, 134, 146, 147, 177, 186, 190, 191, 211, 215, 278, 338, 339, 340, 341, 342, 342, 344, 345, 346, 346, 347, 354, 355, 356, 357, 365, 376, 475, 480, 497  
PDM, 16  
PHP, 63, 339  
Caurulvads, 44, 53, 115, 128, 183, 303, 307, 308, 311, 312, 349, 350, 351, 352, 354, 355, 356, 367, 370, 417, 430, 476  
PLM, 16, 246  
PLN, 8, 272, 296  
Plotly, 320, 336, 337, 337, 372  
PMIS, 3, 24, 32, 62, 152, 196, 198, 211, 239, 240, 245, 246, 247, 248, 249, 250, 272, 326, 351, 387, 388, 389, 390, 458, 462, 487  
PMS, 84, 151, 233  
PostgreSQL, 63, 64, 65, 329, 395  
Power BI, 320, 372, 482  
privātais kapitāls, 464, 487  
Python, 56, 63, 103, 105, 112, 114, 115, 117, 118, 119, 129, 130, 131, 133, 134, 135, 166, 167, 179, 187, 219, 224, 225, 227, 303, 304, 308, 320, 329, 330, 332, 333, 339, 340, 346, 348, 356, 367, 369, 372, 378, 409, 417, 426, 459, 475, 476, 479, 494

## Q

QTO, 72, 196, 214, 215, 216, 217, 218, 219, 221, 223, 225, 226, 228, 237, 238, 242, 243, 251, 301, 302, 475, 479  
QWEN, 103, 104, 124, 129, 133, 166, 219, 221, 300, 329, 349, 343, 349, 381, 429, 443

## R

RAG, 111, 116  
RDBMS, 63, 64, 65, 82, 89.  
RDF, 267, 268, 269  
RegEx, 126, 136, 177, 179, 331, 332, 333, 357, 474  
RFID, 8, 18, 58, 84, 418, 419, 455, 460, 465, 482  
ROI, 311, 317, 319, 321, 351, 370, 389, 479  
RPM, 14, 84, 151, 331, 333, 334  
RTV, 8, 73, 77, 140, 186, 227, 272, 296, 300, 302, 322, 329, 357, 376, 414, 415, 417, 497

## S

SaaS, 24, 50, 51, 52, 458  
 DARBIĀS JOMA, 273, 277  
 SDK, 139, 141, 257, 264, 273, 281, 286, 296, 329, 414  
 Seaborn, 123, 320, 336, 337, 408, 412  
 SPARQL, 269  
 SQL, 63, 65, 66, 88, 89, 103, 105, 119, 123, 166, 168, 268,  
 269, 279, 276, 277, 300, 329, 376, 392, 494  
 SQLite, 63, 64, 65, 166, 167, 296, 329, 363  
 SOLIS, 261, 262, 263, 266, 272, 276, 277, 292  
 Streamlit, 336, 337  
 SVF, 142, 276, 285

## T

Pārveidot, 128, 193, 311, 323, 325, 326, 330, 331, 333, 334,  
 340, 341, 342, 344, 347, 361, 362, 363, 365, 367, 383,  
 384, 450, 475

## U

Vienklase, 154, 155, 156.  
 USD, 142, 143, 255, 276, 277, 278, 279, 280, 281, 284, 285,  
 286, 292, 296

## V

VectorOps, 373, 398, 400, 401  
 VR, 84, 271, 285

## W

WhiteBox, 240, 242, 243.

## X

XLSX, 8, 61, 62, 123, 128, 129, 231, 268, 276, 277, 280, 296,  
 299, 308, 351, 373, 377, 378, 379, 380, 401, 474, 480  
 XML, 61, 88, 89, 92, 128, 143, 269, 272, 277, 280, 280, 291,  
 299, 329, 378, 414, 474